

L'Aspetto Grammaticale come Categoria Musicale

Aktionsart, telicità e il tempo interno degli eventi sonori

Il tempo che non viene misurato

Esiste un modo di parlare del tempo musicale che la teoria non ha ancora trovato. Non il tempo metronomico, che misura la velocità con cui i battiti si succedono e che trova la sua espressione più semplice nella cifra stampata in cima a un brano moderno — quaranta, sessanta, centoventi, duecento alla semiminima. Non il tempo formale, che descrive la successione delle sezioni architettoniche di un brano — esposizione, sviluppo, ripresa, coda in una forma sonata; tema, variazioni, stretta in un'aria con variazioni; A, B, A nella forma tripartita — e che da Koch a Czerny, da Reicha a Marx, ha occupato un posto di assoluta centralità nella teoria musicale di tutti i secoli. Non il tempo armonico, che registra la velocità con cui gli accordi cambiano e le tonalità si alternano, la densità delle progressioni cadenzali, il ritmo delle modulazioni.

C'è un tempo più intimo di tutti questi, un tempo che riguarda non la velocità né la forma né la densità delle progressioni, ma la struttura interna degli eventi musicali come eventi — la loro qualità di essere compiuti o incompiuti, conclusi o aperti, puntualmente istantanei o processualmente distesi nel tempo, diretti verso un termine naturale o espansi senza alcun termine che ne faccia parte. Questo tempo — che la linguistica ha imparato a descrivere con la categoria dell'aspetto grammaticale e che la filosofia del linguaggio ha elaborato con la categoria delle Aktionsart, ovvero delle classi di azione verbale — non ha ancora ricevuto un trattamento sistematico nella teoria musicale, e la sua assenza è la lacuna che questo saggio intende colmare.

La domanda che apre il saggio è tanto semplice nella formulazione quanto inesplorata nelle conseguenze: la musica ha aspetto? Quando una cadenza autentica perfetta raggiunge la sua tonica, c'è una struttura cognitiva analoga a quella del verbo perfettivo russo che descrive un'azione completata — che esprime non soltanto il fatto che l'azione è avvenuta ma che essa si è conclusa in un punto terminale naturale, che ha raggiunto il suo limite interno, che non c'è più nulla da fare per portarla a termine perché il termine è stato raggiunto? Quando un pedale di dominante si distende sotto

un'armonia che oscilla senza cercare risoluzione, c'è una struttura analoga all'imperfettivo russo, al progressivo inglese, alla duratività delle lingue bantu — alla marcatura grammaticale, cioè, di un processo che è in corso senza che vi sia, intrinsecamente, un punto in cui potrebbe "finire" di essere se stesso, un punto in cui l'evento avrebbe esaurito la propria natura e potrebbe cessare di svolgersi? E quando, nel contrappunto a più voci, una parte cadenza mentre un'altra prosegue indifferente su un pedale e una terza percorre una scala di passaggio, le tre voci sono in stati aspettuali diversi e simultanei — una al termine del proprio evento, un'altra nel mezzo di un processo senza termine, la terza nel mezzo di un movimento che convergerà verso un proprio termine distinto?

Queste domande non sono metaforiche nel senso debole del termine. Non si sta dicendo che la musica "assomiglia" al linguaggio in modo vago e intuitivo, che c'è qualcosa nella cadenza che ricorda vagamente la conclusione di un'azione descritta al passato remoto, che c'è qualcosa nel pedale che ricorda vagamente la progressività di un verbo al gerundio. Si sta avanzando la tesi precisa — e difendibile con gli strumenti dell'analisi teorica e dell'evidenza cognitiva — che la struttura cognitiva sottostante alla percezione del tempo interno degli eventi musicali è isomorfa alla struttura cognitiva sottostante alla grammaticalizzazione dell'aspetto nelle lingue naturali, e che questo isomorfismo non è accidentale né puramente metaforico ma rivela qualcosa di fondamentale sull'architettura della mente che processa il tempo in qualsiasi dominio — linguistico, musicale, visivo, motorio.

La musica e il linguaggio condividono, in questa proposta, non un vocabolario né una sintassi superficiale — non si sta dicendo che la musica "ha" soggetti e predicati, né che le progressioni armoniche "sono" frasi nel senso sintattico — ma una categoria cognitiva profonda: la capacità di rappresentare gli eventi non solo come collocati nel tempo ma come strutturati internamente in modi qualitativamente diversi, come aventi o non avendo un punto terminale che fa parte della loro struttura interna, come presentati dall'esterno nella loro interezza o dall'interno nel loro svolgimento. Questa capacità — che si potrebbe chiamare la capacità di vedere gli eventi sotto una certa qualità temporale — è il nucleo di ciò che la linguistica chiama aspetto, e la tesi di questo saggio è che essa non appartiene esclusivamente al linguaggio ma è una delle capacità cognitive generali della mente umana, che si manifesta in forme diverse in domini diversi ma mantiene una struttura profonda comune.

Il percorso argomentativo si sviluppa in cinque capitoli e tre appendici. Il Capitolo I ricostruisce in dettaglio la teoria delle Aktionsart di Vendler e la teoria dell'aspetto grammaticale di Comrie, resistendo alla tentazione di semplificare una distinzione che nella letteratura secondaria è stata spesso fraintesa o appiattita, e integrando questi contributi con quelli di Jackendoff sulla struttura tematica e di Hasty sul metro come processo. Il Capitolo II costruisce la tassonomia degli eventi musicali sulla base delle categorie aspettuali, con analisi ravvicinate di esempi specifici tratti dal repertorio e con attenzione particolare alla forma che il confine tra telico e atelico assume nei diversi contesti stilistici; include una sezione dedicata all'aspetto nella musica vocale

come dominio in cui il livello aspettuale linguistico e quello musicale si sovrappongono, interagiscono e talvolta confliggono in modi cognitivamente ricchi. Il Capitolo III sviluppa la dimensione tipologica: le tradizioni musicali diverse grammaticalizzano l'aspetto in modi sistematicamente diversi, e questa variazione tipologica ha conseguenze cognitive che si estendono all'esperienza dell'ascolto, alla pratica dell'improvvisazione e all'atto della composizione; questo capitolo include la musica elettroacustica e lo spettralismo come tradizione che grammaticalizza l'aspetto attraverso la dimensione timbrica piuttosto che attraverso quella armonica. Il Capitolo IV affronta la diatesi aspettuale nella polifonia, ovvero il fenomeno per cui le diverse voci di un contrappunto possono trovarsi in stati aspettuati simultaneamente diversi, e sviluppa le implicazioni di questa dissincronia per la teoria della tensione musicale e per la comprensione del tempo musicale come tempo plurale e stratificato. Il Capitolo V — la sezione più ampiamente nuova rispetto al nucleo originario di questa ricerca — affronta le tradizioni ritmiche non-occidentali come terreno di verifica della teoria e come fonte di tipologie aspettuati che il solo studio della musica colta occidentale non potrebbe rivelare, con analisi estese del Tala indostanico, del gamelan giavanese e delle polifonie vocali dell'Africa subsahariana. Un'appendice metodologica raccoglie le formalizzazioni in notazione plaintext delle strutture aspettuati analizzate nel testo; una seconda appendice sviluppa il rapporto tra questa teoria e la teoria del metro di Christopher Hasty, mostrando convergenze e divergenze sistematiche; una terza appendice affronta il caso speciale della musica vocale e della stratificazione aspettuale bilingue, approfondendo la relazione tra aspetto grammaticale del testo e aspetto musicale della composizione.

Una nota sullo stile del ragionamento è necessaria prima di procedere. L'aspetto grammaticale è una categoria che divide i linguisti quasi quanto li unisce: le classificazioni di Vendler sono state criticate, modificate, ampliate e contestate in cinquant'anni di letteratura, e non c'è accordo su questioni fondamentali come il confine tra Aktionsart lessicale e aspetto grammaticale, il rapporto tra aspetto e azionalità, o la misura in cui le categorie aspettuati siano universali o culturalmente specifiche. Questo saggio non presuppone che la teoria aspettuale in linguistica sia definitivamente sistemata e quindi esportabile senza problemi al dominio musicale. Presuppone invece che le tensioni interne alla teoria aspettuale siano esse stesse illuminanti quando trasferite alla musica, che le ambiguità e i casi limite aspettuati abbiano controparti musicali precise e analiticamente produttive, e che l'applicazione della teoria alla musica possa gettare nuova luce sia sulla musica sia sulla teoria linguistica, illuminando dall'esterno strutture che dall'interno — cioè dal solo studio delle lingue naturali — non sono visibili nella loro piena generalità cognitiva.

Capitolo I. Le Aktionsart e l'aspetto grammaticale: ricostruzione critica di un framework

1.1 La distinzione fondamentale: tempo esterno e tempo interno

Il filosofo linguista Zeno Vendler, nel suo articolo del 1957 "Verbs and Times" poi ripreso e sistematizzato in *Linguistics in Philosophy* del 1967, parte da un'osservazione che in apparenza potrebbe sembrare ovvia ma che ha conseguenze di straordinaria portata per la teoria del significato verbale e, come si mostrerà, per la teoria degli eventi in qualsiasi dominio cognitivo. I verbi della lingua inglese si comportano diversamente rispetto ai test grammaticali che coinvolgono il tempo, e questa diversità non è casuale né puramente sintattica ma riflette differenze profonde nelle strutture temporali intrinseche degli eventi che i verbi descrivono. "Running" e "writing a letter" sono entrambi processi verbali che occupano tempo — entrambi accettano, ad esempio, la forma progressiva "she is running", "she is writing a letter" — eppure chi interrompe qualcuno nel mezzo di una corsa può dire che quella persona ha corso, mentre chi interrompe qualcuno nel mezzo della scrittura di una lettera non può ancora dire che quella persona ha scritto una lettera. Il processo della corsa è tale che qualunque parte di esso è già, in sé, una corsa; il processo della scrittura della lettera è tale che nessuna parte di esso, presa da sola prima del completamento, è ancora una lettera scritta. Questa differenza non riguarda la velocità, non riguarda la durata, non riguarda la complessità dell'azione: riguarda la struttura interna dell'evento, il suo rapporto con un eventuale termine, la sua logica interna di svolgimento e completamento.

Vendler propone di distinguere quattro classi fondamentali di verbi — che chiama *states*, *activities*, *accomplishments* e *achievements* — sulla base di test come la compatibilità con il progressivo, la risposta alla domanda "per quanto tempo?" versus "in quanto tempo?", e la logica dell'interruzione. Prima di illustrare queste classi con la precisione che il trasferimento al dominio musicale richiede, è però necessario fare una distinzione che Vendler non tematizza esplicitamente ma che la letteratura successiva ha reso centrale e che è fondamentale per evitare confusioni nell'applicazione al dominio musicale: la distinzione tra *Aktionsart* (o azionalità) e aspetto grammaticale.

Le *Aktionsart* — il termine tedesco che significa letteralmente "tipo di azione" — sono proprietà del significato lessicale di un verbo o di una predicazione: appartengono alla semantica del verbo in sé, indipendentemente dai morfemi grammaticali che lo modificano. L'*Aktionsart* di "correre" è quella di un'attività; l'*Aktionsart* di "disegnare un cerchio" è quella di un *accomplishment*; l'*Aktionsart* di "arrivare" è quella di un *achievement*; l'*Aktionsart* di "sapere" è quella di uno stato. Queste sono proprietà che i verbi hanno in virtù del loro significato, prima ancora che qualsiasi grammatica intervenga a modificarli o classificarli.

L'aspetto grammaticale, al contrario, è una categoria grammaticale che si sovrappone all'*Aktionsart* attraverso morfemi dedicati — suffissi, ausiliari, particelle aspettuali — e che può modificare la struttura temporale di un evento indipendentemente dalle sue proprietà lessicali intrinseche. In russo, la distinzione perfettivo/imperfettivo è aspetto grammaticale nel senso più pieno: ogni radice verbale esiste in due forme

morfologicamente distinte, una perfettiva e una imperfettiva, e la scelta tra le due è obbligatoria in qualsiasi contesto di passato — il russo non ammette un passato "non marcato" aspectualmente, non esiste la possibilità di descrivere un'azione passata senza scegliere tra l'una e l'altra forma. In inglese, la distinzione simple/progressive è aspetto grammaticale: "she wrote" versus "she was writing" applica una categorizzazione aspettuale a una medesima radice verbale, presentando lo stesso evento come visto dall'esterno nella sua interezza (simple) o come visto dall'interno nel suo svolgimento (progressive).

Le Aktionsart sono proprietà lessicali che determinano quali aspetti grammaticali sono compatibili con un dato verbo, ma questa compatibilità non è mai assoluta: è un'area di attrazione e resistenza che il contesto può deformare. Il verbo "to know", che è uno state nel senso di Vendler, non accetta normalmente il progressivo — "*she is knowing the answer" è agrammaticale nella gran parte dei contesti — perché la categoria aspettuale del progressivo impone una struttura di processo in corso, di azione che si svolge dinamicamente in un intervallo di tempo, che è cognitivamente incompatibile con la staticità dello stato. Ma questa incompatibilità non è assoluta: in inglese colloquiale, "she is being difficult" usa il progressivo con un verbo di stato per segnalare una qualità momentanea e volitiva piuttosto che permanente — non si sta descrivendo una proprietà stabile di carattere ma un comportamento intenzionale e transitorio. Il che mostra che aspetto grammaticale e Aktionsart interagiscono in modo complesso, che non si limitano a sommarsi meccanicamente, e che il contesto può "forzare" — il termine tecnico è *coercion aspettuale* — un verbo in una categoria aspettuale diversa da quella che la sua Aktionsart lessicale suggerirebbe. Questo fenomeno della coercion aspettuale, che Moens e Steedman (1988) hanno analizzato sistematicamente, è di importanza cruciale per l'analisi musicale, come si mostrerà nel Capitolo IV.

Questa distinzione tra proprietà lessicali e categorie grammaticali è fondamentale per il trasferimento al dominio musicale, e merita di essere sviluppata ulteriormente prima di procedere. Quando si chiede se la musica ha aspetto, si sta chiedendo due cose diverse che vanno tenute separate, pena la confusione sistematica tra livelli di analisi che devono rimanere distinti. Si chiede, in primo luogo, se gli eventi musicali hanno Aktionsart: se la struttura temporale intrinseca di una cadenza autentica è diversa dalla struttura temporale intrinseca di un pedale di tonica, come la struttura temporale intrinseca di "trovare" (achievement) è diversa da quella di "correre" (activity) — se, cioè, a prescindere da qualsiasi grammatica musicale, i diversi tipi di eventi sonori hanno diverse qualità temporali interne che si impongono all'ascolto come distinzioni cognitivamente reali. Si chiede, in secondo luogo, se la musica ha equivalenti funzionali dell'aspetto grammaticale: se esistono strutture musicali — figure retoriche, disposizioni armoniche, pattern ritmici, articolazioni, dinamiche, strutture formali — che modificano la struttura temporale intrinseca di un evento musicale in modo analogo a come un morfema aspettuale modifica la struttura temporale di un verbo, presentando un evento con una data Aktionsart come se avesse una Aktionsart diversa o aggiungendo una prospettiva aspettuale alla sua qualità temporale intrinseca. La

prima domanda è più semplice e la risposta è affermativa con evidenza quasi immediata; la seconda è più difficile e richiede un'analisi attenta delle strutture musicali che fungono da marcatori aspettuali, delle tradizioni che le hanno sviluppate, e dei meccanismi cognitivi attraverso cui producono i loro effetti.

1.2 Le quattro classi di Vendler: una ricostruzione precisa e analitica

Vendler distingue i verbi inglesi in quattro classi sulla base di proprietà che definisce in modo operativo — cioè mediante test linguistici applicabili in modo relativamente meccanico — prima che in modo teorico. Vale la pena ricostruire queste proprietà con una precisione maggiore di quella usuale nella letteratura secondaria, perché il trasferimento al dominio musicale dipende da una comprensione accurata di ciò che ciascuna classe cattura a livello cognitivo, non solo di ciò che ciascuna classe descrive a livello superficiale.

Gli **stati** (*states*) sono predicazioni che descrivono situazioni omogenee, prive di dinamica interna, che non si svolgono nel tempo ma semplicemente valgono in un intervallo di tempo. "Sapere", "credere", "assomigliare", "contenere", "amare" nel senso di attitudine permanente piuttosto che di azione momentanea: queste predicazioni non "avvengono" in senso stretto, non hanno fasi interne che si susseguono e si trasformano, non "fanno" qualcosa nel tempo ma "sono" qualcosa per un certo periodo. Un test operativo classico è la resistenza al progressivo: in inglese, "she is knowing the answer" è agrammaticale o altamente marcato, mentre "she knows the answer" è normale — e la spiegazione cognitiva di questa resistenza è che il progressivo impone la struttura di un processo in svolgimento, che è incompatibile con la natura statica dello stato. Un secondo test è la risposta alla domanda "cosa stava facendo quando l'hai chiamata?": rispondere "stava sapendo la risposta" è strano perché lo stato non è un'attività che si possa "stare facendo" in un dato momento — è una condizione che vale o non vale, non un'azione che si compie progressivamente.

Dal punto di vista della struttura temporale, e questa è la caratteristica più importante per il trasferimento al dominio musicale, gli stati non hanno nemmeno un inizio o una fine che facciano parte della struttura dell'evento stesso. Quando dico "sapevo la risposta da due anni", il mio sapere non ha avuto un inizio strutturalmente necessario né una fine strutturalmente prevista: semplicemente, a un certo punto ho iniziato a saperla — ma questo "inizio" è esterno alla struttura dello stato, è il momento in cui lo stato è subentrato, non un "inizio" nel senso in cui un processo ha un inizio come prima fase del suo svolgimento — e a un certo punto ho smesso di saperla — ma questo "fine" è ugualmente esterno alla struttura dello stato, è il momento in cui lo stato è cessato, non una "conclusione" nel senso in cui un processo ha una conclusione come ultima fase del suo svolgimento. Lo stato è temporalmente "piatto": è uguale a sé stesso in ogni momento della sua durata, non progredisce né si trasforma né tende verso nulla.

Le **attività** (*activities*) sono predicazioni che descrivono processi dinamici, omogenei nel senso specifico e tecnico che qualsiasi parte di essi è un'istanza dello stesso tipo di evento, e privi di un termine naturale intrinseco. "Correre", "nuotare", "spingere il carrello", "leggere" senza un oggetto specifico, "parlare": queste predicazioni descrivono processi che si svolgono nel tempo — hanno dinamica interna, diversamente dagli stati — e tuttavia non tendono verso nessun punto terminale che faccia parte della loro struttura interna. Si può correre per un'ora, per un minuto o per tutta la vita, e in nessuno di questi casi si è "finito di correre" nel senso che l'attività abbia raggiunto il suo termine naturale: ci si ferma, ma il fermarsi è sempre un'interruzione esterna all'attività, una decisione presa dall'esterno della struttura dell'evento, non il suo completamento intrinseco.

Un test operativo è la compatibilità con "per quanto tempo?" in contrapposizione a "in quanto tempo?": "corse per un'ora" è naturale, mentre "raggiunse la vetta per un'ora" è anomalo perché raggiungere la vetta è un achievement puntuale che non si protrae nel tempo. Al contrario, "corse in un'ora" è anomalo nel senso rilevante perché suggerisce un punto di completamento — come "percorse un miglio in un'ora", che aggiunge il terminus di un miglio e trasforma l'attività in accomplishment — mentre "corse per un'ora" è semplicemente descrittivo della durata. Un secondo test cruciale è la logica dell'interruzione: se si interrompe una corsa a metà, si può dire che si è corso — l'attività è avvenuta parzialmente, e la parte svoltasi è già un'istanza dell'evento — mentre se si interrompe il processo di "raggiungere la vetta" prima del termine, non si può dire che si è raggiunta la vetta. Questa differenza nella logica dell'interruzione è la manifestazione linguistica di una differenza cognitiva profonda: le attività hanno la proprietà di essere cumulative e subomogenee (qualsiasi sotto-parte è dello stesso tipo dell'intero), mentre gli accomplishments sono quantizzati e direzionali (nessuna sotto-parte è dello stesso tipo dell'intero, e tutte le sotto-parti sono orientate verso il terminus).

Gli **accomplishments** (*compimenti*) sono predicazioni che descrivono processi dinamici che tendono verso un termine naturale intrinseco e che si completano quando e solo quando quel termine viene raggiunto. "Disegnare un cerchio", "scrivere una lettera", "costruire una casa", "percorrere un miglio", "leggere un libro": queste predicazioni descrivono processi che hanno fasi strutturalmente distinte — un inizio, uno svolgimento orientato, un termine — e il cui completamento richiede non solo che il processo si sia svolto per un certo tempo ma che abbia raggiunto il suo punto terminale definitorio. Un accomplishment interrotto è un accomplishment incompiuto e non semplicemente un accomplishment più breve: se si interrompe il disegno del cerchio prima di chiuderlo, non si è disegnato un cerchio ma un arco, un segmento curvo — un oggetto categorialmente diverso che manca del termine definitorio dell'evento.

La risposta naturale alla domanda "in quanto tempo?" piuttosto che "per quanto tempo?" è un test operativo significativo: "disegnò il cerchio in dieci minuti" è naturale perché descrive il completamento del processo, "corse in dieci minuti" è

anomalo nel senso rilevante perché la corsa non ha un termine naturale che si raggiunge in dieci minuti (a meno di specificare "corse il miglio in dieci minuti", il che aggiunge il terminus di un miglio e trasforma l'attività in accomplishment). La struttura cognitiva degli accomplishments è quella di un vettore: hanno un inizio, un'orientazione verso un terminus, e il terminus stesso come punto di completamento che fa parte della loro definizione.

Gli **achievement** (*raggiungimenti*) sono predicazioni che descrivono eventi istantanei, puntuali, che non hanno durata intrinseca: un cambiamento di stato che avviene in un punto del tempo senza avere fasi intermedie che appartengano alla struttura dell'evento. "Trovare", "riconoscere", "arrivare", "vincere", "morire", "notare", "realizzare", "raggiungere la vetta": queste predicazioni non si svolgono nel tempo ma avvengono in un istante — un punto senza spessore temporale in cui lo stato delle cose cambia irreversibilmente. Il processo che porta all'achievement — il viaggio che precede l'arrivo, la ricerca che precede il trovare — non appartiene alla struttura dell'achievement ma all'accomplishment che lo precede; l'achievement stesso è il punto terminale di quell'accomplishment, il momento della transizione. Un test operativo è la stranezza della domanda "per quanto tempo?": "arrivò alla stazione per dieci minuti" è anomalo perché l'arrivo non si protrae, mentre "arrivò in dieci minuti" è naturale (ma descrive il viaggio, non l'arrivo). Un secondo test è la logica dell'interruzione vista dal lato opposto: se si interrompe il processo "prima che arrivi", la persona non è ancora arrivata — l'achievement non ha una fase parziale che possa essere compiuta in modo incompleto, non esiste un "mezzo arrivare" che corrisponda all'"aver già corso" nel caso dell'attività. Come disse Vendler con la sua consueta eleganza, "if someone wins the race, there is a point when he wins it" — un punto, non un processo.

Queste quattro classi di Vendler sono state criticate, modificate e ampliate da una letteratura vastissima. Moens e Steedman (1988) hanno proposto di sostituire le classi con un sistema più fine basato sulla nozione di "nucleo eventivo" (*event nucleus*) composto da tre componenti strutturali: il processo preparatorio che precede l'evento, la culminazione che costituisce il punto di cambiamento di stato, e lo stato conseguente che persiste dopo la culminazione. Questa struttura tripartita ha il vantaggio di catturare le relazioni interne agli accomplishments in modo più granulare di quanto la classificazione di Vendler consenta, e di spiegare il fenomeno della coercion aspettuale come trasformazione che prende come punto di partenza una delle tre componenti e la espande o la contrae per adattarsi al contesto. Dowty (1979) ha integrato le Aktionsart con la logica temporale e con il sistema di rappresentazione in termini di insiemi di intervalli e di operatori come BECOME e CAUSE, producendo una formalizzazione semantica delle classi vendleriane che permette di derivare le loro proprietà distribuzionali da principi generali più profondi. Pustejovsky (1991) ha proposto una struttura qualia che rappresenta l'Aktionsart come parte della struttura del significato lessicale, articolata in quattro componenti — telica, formale, agentiva, costitutiva — che determinano insieme il comportamento aspettuale delle parole. Brinton (1988) ha esteso il sistema al di là dei verbi, mostrando che anche i

predicati nominali e aggettivali hanno struttura aspettuale nel senso di Vendler, il che suggerisce che le Aktionsart non sono una categoria specificamente verbale ma una categoria eventiva di portata più generale.

Queste raffinzioni non contraddicono la proposta di Vendler ma la elaborano, e per i nostri scopi la tassonomia a quattro classi — integrata con le distinzioni di Comrie tra aspetto perfettivo e imperfettivo e con la nozione di telicità sviluppata da Garey (1957) e da Krifka (1989, 1992) — è sufficientemente precisa e sufficientemente ricca da essere applicata alla musica con produttività analitica.

1.3 Comrie e l'aspetto grammaticale: la prospettiva crosslinguistica

Bernard Comrie, in *Aspect* (1976) — un volume che insieme alla grammatica tipologica di Givón rappresenta il fondamento della prospettiva crosslinguistica nello studio delle categorie verbali — affronta l'aspetto grammaticale non dal punto di vista delle proprietà semantiche dei verbi singoli ma dal punto di vista dei sistemi grammaticali che le lingue hanno sviluppato per codificare le relazioni temporali interne degli eventi. La prospettiva di Comrie è tipologica nel senso preciso che ciò che lo interessa non è la grammatica di una singola lingua — l'inglese, il russo, il cinese — ma la variazione cross-linguistica nelle strategie di grammaticalizzazione dell'aspetto, e le generalizzazioni cognitive che questa variazione rivela sulle capacità percettive e concettuali della mente umana nel trattare il tempo degli eventi. La distinzione fondamentale di Comrie è quella tra aspetto perfettivo e imperfettivo, che opera a un livello diverso e più grammaticale delle Aktionsart di Vendler — non è una proprietà dei verbi come tali ma una scelta che il parlante fa nell'atto di descrivere un evento, una prospettiva che il parlante adotta rispetto all'evento che descrive.

Un evento descritto in aspetto perfettivo è presentato come un tutto — come un'unità coesa che può essere considerata dall'esterno, nella sua interezza, come un oggetto nella sua totalità. Il parlante che usa il perfettivo "si fa un passo indietro" rispetto all'evento e lo vede intero, con il suo inizio, il suo svolgimento e la sua conclusione tutti compresi nell'atto di presentazione. Un evento descritto in aspetto imperfettivo è presentato come un processo — come qualcosa di osservato dall'interno, nel suo svolgimento, mentre ancora si svolge o si svolgeva al momento di riferimento, senza che la prospettiva della presentazione includa né l'inizio né la fine. Il parlante che usa l'imperfettivo "si trova dentro" l'evento e lo descrive da quella prospettiva interna, vedendo il processo ma non la sua cornice temporale.

Comrie chiarisce questa distinzione con un'analogia spaziale che è diventata un punto di riferimento nella letteratura: il perfettivo è come osservare un oggetto dall'esterno, vedendo la sua forma complessiva e i suoi contorni; l'imperfettivo è come osservare lo stesso oggetto dall'interno, vedendo il suo spazio interno, i suoi dettagli, la sua struttura interna, ma senza poter vedere la sua forma esterna. Applicato al tempo: "scrive una lettera" (perfettivo) presenta l'evento come un'unità totale con inizio e fine entrambi inclusi nella presentazione; "stava scrivendo una lettera" (imperfettivo

progressivo) presenta l'evento dall'interno del suo svolgimento, senza includere nella presentazione né l'inizio né la fine, focalizzandosi sul processo in corso al momento di riferimento.

Questa distinzione perfettivo/imperfettivo è grammaticale nel senso preciso e tecnicamente impegnativo che è obbligatoria in certe lingue: in russo, ogni forma verbale al passato è necessariamente o perfettiva o imperfettiva, e la scelta tra le due non è libera ma ha conseguenze semantiche che non possono essere ignorate senza produrre enunciati inaccettabili o semanticamente devianti. "*On napisal pis'mo*" (lui scrisse la lettera, aspetto perfettivo) implica che la lettera è stata completata, che il processo è giunto al suo termine naturale, che c'è una lettera scritta come risultato dell'azione; "*on pisal pis'mo*" (lui stava scrivendo la lettera, aspetto imperfettivo) non implica il completamento — la lettera potrebbe essere rimasta incompiuta, il processo potrebbe essere stato interrotto prima di raggiungere il suo terminus, e l'unico fatto comunicato è che il processo di scrittura era in corso in un certo momento del passato. Questa asimmetria implicativa — la forza della quale è stata dimostrata sperimentalmente da una serie di studi sulla comprensione delle forme aspettuali nel russo e in altre lingue slave — è la caratteristica centrale dell'aspetto perfettivo: implica il raggiungimento del termine naturale dell'evento, ovvero la sua telicità, e questa implicazione è sufficientemente forte da poter essere cancellata solo con un'esplicita negazione contestuale.

Comrie distingue poi, all'interno dell'aspetto imperfettivo, almeno due sottocategorie di importanza particolare per il trasferimento al dominio musicale. Il progressivo codifica un processo in corso al momento di riferimento: descrive l'evento dall'interno del suo svolgimento e implica che al momento indicato il processo non era ancora terminato, che c'era ancora una durata futura del processo che non è compresa nella descrizione. L'abituale codifica la ripetizione regolare di un tipo di evento come proprietà di un agente o di una situazione: non descrive un'azione in corso in un momento specifico ma una regolarità, una disposizione, una tipicità. "Lei leggeva ogni sera" è imperfettivo abituale: non descrive una lettura che era in corso in un momento specifico della sera ma la proprietà di quell'agente di essere un lettore serale, la regolarità di quella pratica.

Questa distinzione tra progressivo e abituale è rilevante per la musica in un modo che non è immediatamente ovvio ma che diventa chiaro non appena si considerano certi tipi di strutture ripetitive: l'ostinato di basso, il passacaglia, la forma del tema con variazioni sono strutture musicali che hanno struttura aspettuale abituale nel senso di Comrie — presentano una tipologia di evento come proprietà di uno spazio musicale piuttosto che come un singolo evento in corso di svolgimento. La distinzione tra il progressivo musicale — un processo in corso che tende verso il suo terminus — e l'abituale musicale — un pattern che si itera come proprietà strutturale dello spazio sonoro — è una delle più sottili e delle più produttive dell'intera teoria.

Un'ulteriore distinzione introdotta da Comrie, di cruciale importanza per il Capitolo IV, è quella tra aspetto risultativo e aspetto esperienziale. L'aspetto risultativo, presente nel perfetto inglese e in molte altre lingue, codifica uno stato presente come risultato di un evento passato: "la porta è aperta" non descrive l'azione di aprire ma lo stato risultante dall'apertura, uno stato che è attuale e percepibile nel momento della descrizione come conseguenza di qualcosa che è avvenuto in passato. L'aspetto esperienziale, distinto nelle lingue che lo grammaticalizzano come il cinese mandarino con la particella *guo*, codifica che un evento è avvenuto almeno una volta nell'esperienza del soggetto, senza implicare che lo stato risultante sia ancora attuale. Entrambe queste categorie hanno controparti musicali che richiedono analisi accurate: la coda di una fuga è, in aspetto risultativo, rispetto allo sviluppo contrappuntistico che l'ha prodotta; la ripresa di un tema trasformato dalla modulazione dello sviluppo ha una qualità esperienziale rispetto alla sua prima presentazione — non è semplicemente il tema che ritorna ma il tema arricchito dall'esperienza dello sviluppo, un'esperienza che l'ascoltatore porta con sé nell'ascolto della ripresa come bagaglio cognitivo che colora la nuova occorrenza.

1.4 Telicità come proprietà centrale: la struttura cognitiva dell'endpoint

La nozione di telicità — dal greco *telos*, fine, termine, scopo — è, tra tutte le proprietà aspettuali discusse nella letteratura, quella di maggiore importanza per il trasferimento al dominio musicale, e merita un'analisi che vada ben oltre la semplice definizione dizionaristica. La telicità è la proprietà degli eventi che hanno un termine naturale intrinseco: un punto in cui l'evento si completa e cessa di essere se stesso perché ha raggiunto il suo punto definitorio, il punto che appartiene alla sua struttura come la punta appartiene alla freccia. Gli accomplishments e gli achievements sono telici; gli stati e le attività sono atelici.

La telicità non è una proprietà della velocità né della durata di un evento: un accomplishment può essere veloce o lento, breve o lungo, senza per questo cambiare la sua natura di processo orientato verso un terminus. La telicità non è nemmeno una proprietà della presenza di una fine: ogni evento finisce, ma non ogni evento ha una fine naturale intrinseca nel senso rilevante. Un'attività finisce quando ci si ferma, quando una decisione esterna all'attività stessa la interrompe, ma non perché l'attività abbia "esaurito" la propria natura e non potrebbe continuare ulteriormente senza cambiare tipo. Un accomplishment finisce quando raggiunge il suo terminus, e questo terminus non è un'interruzione arbitraria ma il punto in cui l'evento si completa, il punto che l'evento stesso "cerca" nel corso del suo svolgimento. La differenza cognitiva tra i due è la differenza tra fermarsi e arrivare: ci si ferma dall'esterno, si arriva dall'interno della struttura dell'evento.

La struttura formale della telicità è stata elaborata con precisione da Manfred Krifka in una serie di contributi degli anni Ottanta e Novanta. In termini molto schematici, Krifka mostra che la distinzione telico/atelico corrisponde alla distinzione tra eventi

quantizzati ed eventi cumulativi in una semantica degli eventi formale. Un evento atelico come una corsa è cumulativo nel senso che la somma di due corse è ancora una corsa — la proprietà si chiude sotto la composizione di eventi dello stesso tipo. Un evento telico come "scrivere una lettera" non è cumulativo in questo senso: la somma di due eventi di "scrivere una lettera" non è un singolo evento di "scrivere una lettera" ma due eventi separati, ciascuno con il proprio terminus. L'intuizione formale di Krifka è che la telicità è la proprietà degli eventi che hanno un "confine destro" naturale nella struttura eventiva — un punto massimale al di là del quale l'evento non può estendersi senza cambiare tipo — mentre l'atelicità è la proprietà degli eventi per cui non esiste tale confine destro, e l'evento potrebbe sempre estendersi ulteriormente di un qualsiasi ϵ senza cambiare tipo. Questa caratterizzazione formale è ripresa e formalizzata nell'Appendice A, dove viene applicata sistematicamente agli eventi musicali.

Applicata alla musica, la nozione di telicità permette di distinguere con precisione categorie di eventi che la teoria tradizionale tende a collocare nella stessa classe o a distinguere solo sulla base di criteri formali superficiali. Una cadenza autentica perfetta è telica nel senso più preciso: ha un terminus definitorio — la tonica in posizione fondamentale, alla voce superiore e al basso — che completa l'evento e lo distingue categorialmente da una cadenza sospesa (che non raggiunge il terminus) o da una cadenza d'inganno (che sostituisce il terminus atteso con un terminus diverso e inatteso, producendo un achievement cognitivo di reinterpretazione). Un ostinato di basso è atelico nel senso preciso: il basso potrebbe sempre ripetersi un'altra volta, una volta ancora, e ancora, senza che la struttura cambi tipo, e il suo "fine" è sempre e necessariamente una decisione compositiva esterna alla struttura dell'evento, non un completamento interno. Una singola nota tenuta senza ornamentazione è stativa: non ha dinamica interna, non tende verso nessun punto, non si svolge nel tempo ma semplicemente sussiste in un intervallo temporale, uguale a sé stessa in ogni momento della sua durata. Un'appoggiatura con risoluzione è un achievement: l'evento rilevante — la risoluzione della dissonanza sulla consonanza — è strutturalmente puntuale, avviene in un istante senza fasi preparatorie che appartengano alla struttura dell'achievement stesso.

1.5 Il contributo di Jackendoff: struttura tematica e struttura eventiva

Ray Jackendoff, in una serie di contributi tra gli anni Ottanta e Novanta culminati in *Semantic Structures* (1990) e nell'articolo "Musical Parsing and Musical Affect" (1991), ha proposto che la struttura concettuale del significato linguistico sia organizzata secondo funzioni semantiche ricorrenti — pattern cognitivi profondi — che si applicano a domini diversi: azione, spazio, possesso, attributo, informazione. La struttura di un predicato come "muoversi da A a B" è una struttura concettuale di tipo PATH applicata al dominio spaziale, ma la stessa struttura concettuale di tipo PATH — con le sue componenti di sorgente, traiettoria e meta — può essere applicata al dominio temporale ("dall'alba al tramonto"), al dominio del possesso ("passare la proprietà da

X a Y"), al dominio attributivo ("diventare da freddo a caldo"), al dominio informativo ("comunicare un'informazione da un mittente a un destinatario").

Questo parallelismo strutturale tra domini — che Jackendoff chiama "isomorfismo tematico" e che ha elaborato nella teoria delle funzioni concettuali come GO, BE, CAUSE, CHANGE applicate a diversi "tipi semantici" di argomenti — suggerisce che la struttura eventiva non è una proprietà esclusiva di nessun dominio particolare ma un modulo cognitivo applicabile universalmente, la cui manifestazione varia a seconda del dominio in cui opera ma la cui struttura di base rimane costante. Per i nostri scopi, il contributo di Jackendoff è rilevante precisamente perché mostra che la struttura aspettuale degli eventi — la struttura di evento come processo con o senza terminus, come stato o come cambiamento — è un aspetto della struttura concettuale generale, non una proprietà specifica del linguaggio verbale in quanto sistema semiotico. Se la struttura "GO(x, FROM_a TO_b)" è applicabile a eventi spaziali, temporali, possessivi e attributivi, non c'è nessun motivo a priori per escludere che sia applicabile anche a eventi musicali, con la differenza che il "dominio" su cui opera è lo spazio tonale — o lo spazio timbrico, o lo spazio dell'intensità dinamica — piuttosto che lo spazio fisico o il tempo cronologico.

Il risultato jackendoffiano più direttamente utile per questa monografia è la proposta che la struttura tematica degli eventi ha due livelli cognitivi distinti: il livello eventivo, descritto da funzioni che codificano cambiamenti nel tempo (GO, CAUSE, CHANGE, MOVE), e il livello di stato, descritto da funzioni che codificano situazioni stabili (BE, STAY, HAVE). Le Aktionsart emergono dalla combinazione specifica di questi livelli e delle loro proprietà: uno stato puro (BE senza dinamica) è stativo; un'attività (GO senza endpoint) è atelica; un accomplishment (GO con endpoint) è telico; un achievement (CHANGE puntuale) è telico e privo di durata interna. Questo sistema fornisce una base teorica per la proposta che gli eventi musicali abbiano struttura aspettuale: hanno struttura eventiva nel senso jackendoffiano, applicata al dominio del suono e organizzata secondo le stesse funzioni concettuali che organizzano la struttura eventiva nel dominio del linguaggio, dello spazio, del possesso.

Un aspetto particolarmente produttivo della proposta jackendoffiana, per l'analisi musicale, è la distinzione tra struttura tematica di primo livello e struttura tematica di secondo livello. La struttura di primo livello descrive l'evento nella sua globalità: questo è un accomplishment, questo è un'attività, questo è uno stato. La struttura di secondo livello descrive le relazioni interne tra i componenti dell'evento: in un accomplishment, le relazioni tra la fase preparatoria, la culminazione e lo stato conseguente; in un'attività, le relazioni di omogeneità tra le parti successive; in uno stato, le relazioni di identità tra ogni momento e il resto. Questa stratificazione è cruciale per analizzare gli eventi musicali di grande scala — come una forma-sonata o un ciclo di variazioni — che hanno una struttura aspettuale di primo livello e strutture aspetuali di secondo livello per i loro componenti, e queste strutture di livelli diversi

interagiscono in modi che richiedono l'analisi jackendoffiana per essere descritti con precisione.

1.6 Hasty e il metro come processo: la dimensione aspettuale del ritmo

Christopher Hasty, in *Meter as Rhythm* (1997) — un'opera impegnativa, densa e deliberatamente lontana dall'ortodossia della teoria musicale nordamericana — propone una teoria del metro musicale che è in larga misura, anche se Hasty non usa mai questo termine, una teoria dell'aspetto ritmico. La tesi centrale di Hasty è che il metro non è una griglia astratta di unità uguali sovrimpressa al ritmo dall'esterno — come se il metro esistesse indipendentemente dal suono e il suono dovesse poi adattarsi a questa griglia — ma è esso stesso un processo di proiezione emergente dalla struttura temporale degli eventi ritmici: ogni evento ritmico è, nella percezione di chi ascolta, un atto di "proiettare" la durata dell'evento appena iniziato come durata attesa per l'evento successivo, e questa proiezione è il metro stesso, non un suo riflesso o una sua conseguenza.

La rilevanza di Hasty per la teoria dell'aspetto musicale è profonda e si estende in diverse direzioni. La sua nozione di "becoming" — il fatto che ogni evento musicale è sempre in un processo di diventare, di aprire una durata prospettica nel momento stesso in cui inizia — è precisamente la struttura cognitiva che nella terminologia aspettuale distingue un evento in corso da uno stato: un evento che "inizia" nel senso di Hasty è un evento che proietta una durata futura, che apre uno spazio temporale che attende di essere riempito o chiuso; uno stato non ha questo tipo di inizio, è già presente come condizione che vale, senza che ci sia un momento di "apertura" che proietti una durata futura da completare. Questa distinzione tra "iniziare-come-proiettare" e "essere-già-presente" — tra il dinamismo di un inizio che apre il futuro e la staticità di una presenza che già è quello che è — è esattamente la distinzione cognitiva tra evento e stato nelle categorie di Vendler, espressa nel linguaggio di una teoria del ritmo piuttosto che in quello di una semantica verbale.

La nozione di "anacrusis" in Hasty — il prefisso metrico che anticipa e prepara un evento culminante, che "appartiene" al battere forte verso cui si dirige — ha la struttura esatta di un accomplishment incompleto: è la fase preparatoria di un evento telico, una fase che si svolge nel tempo verso un terminus che fa parte della struttura dell'evento complessivo, e che acquisisce il suo significato e la sua funzione solo in riferimento a quel terminus. Un'anacrusis che non arriva alla sua downbeat — interrotta prima di raggiungere il terminus metrico — è un'anacrusis incompleta, esattamente come un cerchio incompiuto non è un cerchio ma un arco. La "downbeat" — il tempo forte su cui un evento ritmico culmina — è la culminazione dell'accomplishment anacrusitico: il punto in cui il processo preparatorio raggiunge il suo endpoint, il terminus verso cui tutta l'anacrusis era orientata. Il "continuation" — la continuazione che segue la downbeat — è la fase conseguente nel senso di Moens e

Steedman: il risultato stabile che segue il raggiungimento del terminus, lo stato post-achievement in cui l'evento si è assestato dopo la culminazione.

Questa struttura hastyiana — anacrusis, downbeat, continuation — è isomorfa alla struttura eventiva degli accomplishments di Vendler: preparazione, culminazione, stato conseguente. Hasty non usa mai il termine "aspetto" in questo contesto, e non si collega esplicitamente alla letteratura vendleriana o a quella di Comrie; ma la sua teoria è, tra tutte le teorie musicali disponibili, quella che più si avvicina a una teoria dell'aspetto musicale, e questo saggio la usa come base per sviluppare la proposta con maggiore sistematicità e rigore formale. Va tuttavia notata una lacuna importante nella teoria di Hasty: essa è, nonostante la sua sofisticazione, fondamentalmente monodica nella prospettiva. Hasty analizza come un singolo filo temporale si struttura aspettualmente, non come fili temporali multipli e simultanei interagiscono nelle loro diverse strutture aspetttuali. Questa limitazione — che non inficia la validità della teoria nel suo dominio — è precisamente ciò che il Capitolo IV di questa monografia si propone di superare sviluppando una teoria della diatesi aspettuale nella polifonia.

1.7 Il problema del trasferimento: rischi metodologici e strategie di risposta

Prima di procedere all'applicazione delle categorie aspetttuali alla musica, è necessario affrontare con onestà una difficoltà metodologica che la ricchezza del framework finora esposto rischia di oscurare. Il trasferimento delle categorie linguistiche a un dominio non linguistico è sempre un'operazione rischiosa, e il rischio specifico in questo caso è di produrre un'analogia attraente ma superficiale — un'analogia che usa le parole della teoria aspettuale senza catturarne la struttura operativa, che etichetta gli eventi musicali con le categorie di Vendler senza fornire criteri per verificare se le etichette siano appropriate o arbitrarie. La storia dell'applicazione di categorie linguistiche alla musica — dalla "sintassi musicale" di Riemann al "semiotismo" di Nattiez, dal "narrativismo" di Tarasti alla "topic theory" di Ratner e Allanbrook — è costellata di casi in cui la terminologia linguistica è servita da ornamento retorico piuttosto che da strumento analitico, in cui la citazione di Saussure o di Greimas ha dato alla teoria musicale un'aria di rigore senza un corrispondente aumento di precisione descrittiva o di potere predittivo.

Il rischio si materializza in questo saggio in due forme distinte. Il primo è quello dell'etichettatura arbitraria: applicare le etichette "telico", "atelico", "stativo", "achievement" agli eventi musicali senza un criterio operativo che permetta di distinguere i casi dubbi, di risolvere i casi di ambiguità, di falsificare le classificazioni proposte nel senso popperiano. In linguistica, le categorie aspetttuali si applicano mediante test grammaticali precisi che, pur non essendo infallibili, hanno il vantaggio di essere intersoggettivamente verificabili. In musica, non ci sono morfemi aspetttuali che possano essere applicati meccanicamente: i criteri operazionali devono essere costruiti ad hoc, a partire dalle proprietà degli eventi musicali stessi, e la loro

applicazione richiede un livello di interpretazione teorica che la linguistica operativa non richiede.

Il secondo rischio è quello dell'universalismo inconsapevole: assumere che le categorie aspettuali siano universali — che tutti gli eventi musicali di tutte le tradizioni abbiano aspetto nel senso elaborato da Vendler e Comrie — quando in realtà la grammaticalizzazione dell'aspetto è tipologicamente variabile anche nelle lingue naturali. Ci sono lingue senza morfologia aspettuale obbligatoria, lingue in cui le distinzioni temporali interne degli eventi sono lasciate implicitamente al contesto senza essere codificate grammaticalmente, lingue in cui le categorie corrispondenti a quelle di Vendler sono ritagliate in modo diverso. Se la grammaticalizzazione dell'aspetto è tipologicamente variabile nelle lingue naturali, le tradizioni musicali potrebbero variare analogamente, e un'applicazione acritica delle categorie vendleriane alla musica di tutte le tradizioni potrebbe imporre un modello cognitivo specifico di una tradizione (quella tonale-armonica occidentale) come se fosse universale.

La risposta al primo rischio consiste nello sviluppare criteri operazionali specifici per le categorie aspettuali musicali — criteri analoghi, per struttura e per funzione epistemica, ai test linguistici di Vendler e Comrie, anche se necessariamente diversi nei meccanismi concreti — e di applicarli con la massima coerenza possibile alle analisi dei capitoli successivi. Quattro criteri di questo tipo sono proposti e sistematizzati nel Capitolo II, dove vengono applicati a una gamma di esempi abbastanza varia da mostrare sia la produttività del framework sia i suoi limiti. La risposta al secondo rischio è di adottare fin dall'inizio un'impostazione tipologica che non assume l'universalità delle categorie ma la indaga come una questione aperta. Il Capitolo III sviluppa questa tipologia in modo sistematico, il Capitolo V la estende alle tradizioni non-occidentali, e le Conclusioni riflettono sulle implicazioni teoriche di una tipologia aspettuale musicale che sia genuinamente crossculturale.

2. La Tassonomia degli Eventi Musicali

Telicità, atelicità, statività e achievement nella struttura sonora

2.1 Costruire i criteri operazionali: come si riconosce un evento telico in musica

Poiché la musica non ha morfemi aspettuali espliciti — non esiste un suffisso che trasformi un evento musicale da atelico a telico come il prefisso del perfetto russo trasforma una radice imperfettiva in perfetta, non esiste un ausiliare che segnali la

progressività di un processo musicale come il verbo "essere" con il gerundio segnala la progressività in italiano — i criteri operazionali per identificare l'Aktionsart di un evento musicale devono essere costruiti a partire dalle proprietà intrinseche degli eventi stessi, attraverso un ragionamento analogico che prenda come modello la struttura dei test linguistici senza copiarne meccanicamente la forma. Propongo qui quattro criteri, analoghi per struttura e per funzione epistemica ai test linguistici di Vendler e Comrie anche se diversi nei meccanismi concreti, e li applico sistematicamente nelle analisi dei paragrafi successivi di questo capitolo.

Prima di esporli, è opportuno chiarire il loro statuto metodologico: non si tratta di algoritmi di classificazione che producono risultati univoci e incontrovertibili, ma di strumenti euristici che orientano l'analisi verso strutture cognitivamente rilevanti e che, applicati con coerenza, producono descrizioni verificabili e confrontabili. Nella misura in cui la loro applicazione genera disaccordi — e alcune di queste applicazioni ne genereranno certamente — il disaccordo stesso è analiticamente produttivo: rivela le strutture di ambiguità aspettuali che corrispondono, nel dominio musicale, alle zone grigie che la letteratura linguistica ha ampiamente documentato nel dominio delle lingue naturali.

Il **primo criterio** è quello della sospendibilità senza incoerenza: un evento musicale è atelico se può essere interrotto in qualsiasi momento senza che la struttura dell'evento sia compromessa, ovvero senza che la parte svoltasi fino al punto di interruzione risulti "incompleta" nel senso di aver mancato un terminus naturale che faceva parte della struttura dell'evento. Questo criterio è l'equivalente musicale del test di interruzione di Vendler: se si interrompe una corsa a metà, si può dire che si è corso, ma se si interrompe il disegno del cerchio prima di chiuderlo, non si è disegnato un cerchio. Applicato alla musica: un ostinato di basso è sospendibile in qualsiasi momento, perché qualsiasi numero intero di ripetizioni del pattern costituisce un'istanza completa e strutturalmente coerente dell'evento ostinato, senza che nessun numero di ripetizioni sia intrinsecamente più "completo" di un altro. Una cadenza autentica, al contrario, è non-sospendibile senza incoerenza strutturale: se si interrompe la progressione V7-I prima della risoluzione sulla tonica, l'ascoltatore percepisce non un evento completo di un certo tipo ma un evento categorialmente incompleto — non una cadenza diversa ma una cadenza mancata, una struttura che ha aperto un processo e non lo ha chiuso. L'incoerenza che si produce non è di tipo stilistico o idiomatico ma di tipo strutturale-cognitivo: qualunque ascoltatore con esperienza del sistema tonale percepirà l'incompletezza come tale, indipendentemente dalle sue preferenze estetiche.

Il **secondo criterio** è quello della cumulatività: un evento musicale è atelico se due sue istanze consecutive producono un'istanza dello stesso tipo di evento — se, cioè, la "somma" di due occorrenze dell'evento è ancora un'occorrenza dello stesso evento, solo più lunga. Questo criterio corrisponde alla proprietà della "subinterval property" delle attività formalizzata da Krifka: qualsiasi sotto-intervallo di un'attività è ancora un'istanza della stessa attività, e qualsiasi concatenazione di attività dello stesso tipo è

ancora un'istanza della stessa attività. Due ripetizioni di un ostinato sono ancora un ostinato — un ostinato più lungo, con più ripetizioni, ma strutturalmente dello stesso tipo. Due cadenze autentiche consecutive non sono "una cadenza più lunga" ma due cadenze distinte e separate: la prima ha raggiunto il suo terminus e si è conclusa come evento autonomo, la seconda è un nuovo evento dello stesso tipo che inizia dopo la conclusione del primo. Questa non-cumulatività delle cadenze è la manifestazione musicale della non-cumulatività degli accomplishments nel senso di Krifka.

Il **terzo criterio** è quello della direzionalità interna: un evento musicale è telico se ha una struttura interna orientata verso un terminus — se le sue parti "puntano" verso un endpoint che fa parte della struttura dell'evento, non come obiettivo esterno ma come componente definitoria della struttura stessa. Una progressione di accordi ii-V-I è direzionale nel senso preciso e analiticamente verificabile che il secondo grado "punta" verso il quinto e il quinto "punta" verso il primo: ogni accordo ha una funzione che si definisce rispetto all'accordo successivo e, in ultima istanza, rispetto alla tonica come endpoint del processo cadenzale. Questa direzionalità non è arbitrariamente imposta dall'analista ma è percepita dall'ascoltatore come proprietà strutturale della progressione: nessuno che abbia esperienza del sistema tonale percepirebbe la progressione ii-V come "già completa" — la si percepisce come orientata verso qualcosa che ancora non è avvenuto, come un vettore che punta verso la tonica con una forza proporzionale alla distanza da essa e alla densità delle aspettative armoniche create. Un'armonia modale stabile, al contrario — come un drone di tonica nella musica medievale o nella modalità nordindiana — non ha questa direzionalità interna: ogni momento dell'armonia è ugualmente "a destinazione", ugualmente stabile, ugualmente privo di una tensione che lo spinga verso qualcos'altro.

Il **quarto criterio** è quello dell'irriducibilità dell'endpoint: un evento musicale telico ha un terminus che non è intercambiabile con nessun altro senza distruggere l'identità categoriale dell'evento. La cadenza autentica perfetta ha come terminus la tonica in posizione fondamentale alla voce superiore e al basso, e questo terminus è definitorio nel senso che la sostituzione di qualsiasi suo elemento produce un evento categorialmente diverso: sostituire la tonica con la dominante produce una cadenza sospesa, che è un tipo di evento con un diverso terminus e una diversa funzione strutturale; sostituire la tonica al basso con la tonica al tenore produce una cadenza autentica imperfetta; sostituire la tonica con il sesto grado produce la cadenza d'inganno, il cui effetto percettivo — la sorpresa, la delusione dell'aspettativa soddisfatta — dipende precisamente dal fatto che il terminus atteso viene sostituito con un terminus inatteso, che l'irriducibilità del terminus viene violata. Questa irriducibilità del terminus è la proprietà che distingue la telicità dall'atelicità nel senso più profondo: l'evento telico è definito dal suo terminus specifico, mentre l'evento atelico non ha bersaglio definitorio e la sua conclusione è sempre contingente, sempre esterna alla sua struttura.

2.2 Gli stati musicali: la grammatica dell'immobilità e la sua ricchezza cognitiva

Gli stati musicali — le situazioni sonore che "valgono" in un intervallo di tempo senza svolgersi dinamicamente verso nessun terminus — sono la categoria aspettuale musicale più facile da riconoscere nella sua forma prototipica e al tempo stesso la più difficile da analizzare in modo non triviale. Sembrano non essere eventi in senso pieno, sembrano essere la categoria residuale di "tutto ciò che non fa nulla", e tuttavia la loro presenza o assenza, la loro durata e la loro posizione strutturale determinano in modo radicale il carattere temporale di un brano o di una sezione, la qualità dell'esperienza del tempo che l'ascoltatore vive in quei momenti.

Il caso prototipico di stato musicale è il pedale di tonica: una nota grave tenuta sulla tonica, o un accordo di tonica ripetuto senza progressione verso altri gradi, mentre sopra di esso si svolgono eventi armonici di vario tipo. Il pedale di tonica non "fa" niente nel senso in cui una progressione cadenzale "fa" qualcosa: non si muove, non tende verso nessun terminus, non accumula tensione che attende di essere scaricata, non apre un processo che attende di essere chiuso. È un'armonia che è già nel posto in cui deve essere, che non ha bisogno di andare da nessuna parte perché è già arrivata, che "vale" semplicemente come proprietà del momento sonoro senza che questo valere implichi alcuna dinamica temporale. Ma questa sua staticità non è un'assenza di contenuto cognitivo — è, al contrario, una struttura cognitiva positiva e specifica: è la struttura di ciò che Vendler chiama stato, la struttura di una predicazione che "vale" in un intervallo senza avere la dinamica di un processo.

La prova più diretta che gli stati musicali sono strutture cognitive positive e non semplici assenze di evento viene dalla teoria della tensione armonica di Lerdahl e Jackendoff (1983) e dalla successiva e più sistematica teoria dello spazio tonale di Lerdahl (2001). In questi sistemi, rigorosi nella loro formalizzazione e solidamente supportati empiricamente, la tensione armonica è definita e calcolata come distanza nello spazio tonale tra l'accordo corrente e la tonica di riferimento: un accordo di tonica ha tensione zero, nel senso preciso che la sua distanza dalla tonica è nulla per definizione; un accordo di settima di dominante ha tensione massimale all'interno del sistema diatonico. Uno stato musicale di tensione zero — un accordo di tonica prolungato, un pedale di tonica che si estende per più misure — non è "niente": è la condizione di massima stabilità del sistema, il punto in cui tutte le tensioni sono risolte e l'ascoltatore si trova nella posizione armonico-cognitiva che il sistema tonale considera il suo centro gravitazionale, il suo posto di arrivo.

Ma gli stati musicali non si limitano al pedale di tonica, né sono una prerogativa del sistema tonale. Il drone della bordura medievale — la quinta o l'ottava bassa tenuta sotto le voci superiori nelle organa di Notre Dame o nei canti di Hildegarda di Bingen — produce uno stato musicale di tipo diverso: non è l'arrivo alla tonica di un sistema gerarchico funzionale ma la presenza di un suono-radice che definisce lo spazio modale senza implicare nessuna direzione all'interno di quello spazio. Il pedale di quinta dei corni nella musica del classicismo viennese — le celebri "quinte di corno" che aprono o chiudono le esposizioni delle sinfonie di Haydn e Mozart — è uno stato di colore orchestrale che dipinge un carattere statico e bucolico senza generare alcuna

tensione funzionale verso un terminus. La nota tenuta del muezzin nell'*adhan* islamico è uno stato contemplativo che "è" piuttosto che "va verso". La bordura di tanpura nella musica classica nordindiana — il bordone di quattro note ripetute che scandisce ogni esecuzione di raga — è uno stato armonico-timbrico che non ha dinamica interna né direzionalità ma crea lo sfondo stabile su cui gli eventi del raga si proiettano. Tutti questi dispositivi, provenienti da tradizioni lontanissime e da sistemi sonori eterogenei, producono stati musicali nel senso aspettuale del termine, e questa struttura cognitiva condivisa è la prova che la categoria dello stato musicale non è una specificità del sistema tonale ma un tipo di evento cognitivo di portata universale.

Una delle obiezioni più frequenti a questa categorizzazione degli stati musicali merita di essere sviluppata con cura prima di poter essere risposta, perché la sua solidità è reale. L'obiezione è la seguente: ciò che si chiama "stato musicale" non è affatto privo di dinamica interna al livello della struttura di superficie. Un pedale di tonica in Bach non è silenzio: ci sono note che si muovono sopra di esso, ritmi che pulsano, ornamentazioni che si svolgono, voci che si intrecciano e si rispondono. Un bordone di tanpura non è un suono fisso e immobile: ha una vibrazione, un battimento, una ricchezza timbrica che cambia nel tempo. In che senso si può dire che questi eventi sono "statici" quando il tessuto sonoro che li costituisce o li accompagna è dinamico?

Questa obiezione è corretta nell'identificare un problema reale di livelli di analisi, ma trae da esso una conclusione troppo rapida. La distinzione tra stato e processo negli eventi musicali non riguarda la presenza o l'assenza di movimento nel tessuto sonoro di superficie: riguarda la struttura temporale del livello armonico-funzionale, che è il livello rilevante per l'analisi aspettuale degli eventi musicali come eventi. Un pedale di tonica in una fuga bachiana può avere, alla voce superiore, un tema che corre rapidamente con crome e semicrome, ornamentazioni cromatiche, salti ampi e ritmi puntati: ma la struttura armonica complessiva è in uno stato di tonica che non tende verso nessun altro termine. Il flusso ritmico di superficie è dinamico, ma la struttura funzionale profonda è stativa. È come dire che "lei sa la risposta" non descrive una situazione statica in assoluto — biologicamente, lei sta respirando, i suoi neuroni stanno sparando, il suo cuore sta battendo — ma descrive una situazione stativa al livello del predicato rilevante (il sapere), che non ha dinamica interna indipendentemente da ciò che succede a livello subpersonale. La statività è una proprietà del livello di descrizione rilevante, non dell'oggetto nella sua totalità fisica, e il livello rilevante per l'analisi aspettuale musicale è quello della struttura funzionale degli eventi armonici.

La fenomenologia dell'ascolto offre un'ulteriore conferma della realtà cognitiva degli stati musicali. La percezione del tempo in una sezione stativa è radicalmente diversa dalla percezione del tempo in una sezione di progressione cadenzale. Quando si ascolta un lungo pedale di tonica — le battute di apertura del primo preludio del *Clavicembalo Ben Temperato*, o le misure conclusive di molti corali di Bach in cui la tonica si protrae dopo la cadenza finale — il tempo si comporta diversamente: non "corre" verso qualcosa, non crea aspettative di risoluzione che attendono di essere soddisfatte,

non genera quella qualità di tensione-anticipatoria che caratterizza le sezioni di progressione cadenzale. Questa differenza nella qualità del tempo percepito è il correlato fenomenologico della distinzione aspettuale tra stato e processo, e la sua realtà cognitiva è confermata da studi sulla percezione del tempo musicale che mostrano come le sezioni armonicamente stabili vengano percepite come più "lente" rispetto alle sezioni armonicamente dinamiche, anche a parità di tempo metronomico. Il tempo si "dilata" soggettivamente in uno stato musicale perché il meccanismo cognitivo che misura il tempo attraverso il numero di eventi attenzionalmente rilevanti viene privato del suo input principale: non ci sono progressioni verso cui orientarsi, non ci sono aspettative cadenzali che creino "ancore" temporali, non ci sono eventi telici che delimitino il flusso temporale con i loro terminus.

2.3 Le attività musicali: il processo senza termine e la doppia stratificazione aspettuale

Le attività musicali sono eventi che si svolgono dinamicamente nel tempo — non sono stativi nel senso in cui lo è un pedale di tonica, hanno movimento interno, hanno qualità sonora che cambia istante per istante — ma che non hanno un terminus intrinseco, un endpoint che faccia parte della loro struttura e che sarebbe "raggiunto" quando l'evento si completa. Il caso prototipico è il *sequence* — la sequenza accordale — ovvero la progressione di accordi per movimenti regolari e ripetuti: discesa per terze, discesa per quinta, ascesa cromatica di accordi di settima, discesa per gradi della scala con alternanza di accordi di tonica e di dominante. Il sequence si ripete: ogni ripetizione del pattern aggiunge al processo senza completarlo, porta avanti il movimento senza avvicinarsi a un terminus specifico che faccia parte della struttura del processo stesso, e la fine della sequenza è sempre e inevitabilmente una decisione compositiva esterna alla struttura del processo stesso.

La categoria delle attività musicali richiede, nel contesto specifico del sistema tonale, una precisazione importante che le analisi tradizionali tendono a lasciare implicita e che costituisce forse la distinzione teorica più sottile di questo intero capitolo: la distinzione tra la struttura aspettuale di superficie e la struttura aspettuale di profondità, e la possibilità che le due siano diverse — che un evento sia atelico nel suo svolgimento locale di superficie e al tempo stesso orientato teleologicamente a un livello armonico profondo verso un terminus che la superficie non rende visibile ma che la struttura tonale implica con forza crescente man mano che il processo avanza. Questa distinzione, che chiamo *struttura di telicità proiettata*, è necessaria per capire correttamente la struttura aspettuale del primo preludio del *Clavicembalo Ben Temperato* di Bach (BWV 846) — il brano che più frequentemente viene citato come esempio di attività musicale "pura" — e per capire un meccanismo aspettuale pervasivo nella musica tonale.

Il Preludio in Do maggiore è costruito su un principio di estrema semplicità: ogni misura arpeggia un accordo, e le misure si succedono in una progressione armonica che attraversa diversi gradi della tonalità. A livello di superficie, il brano sembra il caso

paradigmatico di un'attività atelica: ogni misura è un'arpeggiatura dell'accordo del momento, le misure si susseguono in modo omogeneo — ogni misura è dello stesso tipo di ogni altra — e non c'è nessuna ragione intrinseca alla struttura di superficie per cui la misura diciannove debba essere diversa dalla misura venti o dalla misura venticinque. Se il brano si fermasse alla misura undici, o alla quindici, o alla ventidue, la struttura di superficie non mostrerebbe nessuna "incompletezza" nel senso rilevante: si è arpeggiato un certo numero di accordi, e il processo di arpeggio è formalmente completo in qualsiasi momento. Questo è il test della sospendibilità senza incoerenza: l'interruzione in qualsiasi momento produce una struttura non-incompleta.

Ma questa lettura, pur corretta nel descrivere la superficie, ignora qualcosa di cognitivamente reale e percettivamente significativo: a livello armonico profondo, il Preludio percorre una traiettoria inequivocabilmente direzionata che ha la struttura di un accomplishment. Le misure iniziali stabiliscono la tonalità di do maggiore attraverso una progressione che si allontana progressivamente dalla tonica — una progressione che ha la struttura di un accomplishment armonico di ampio respiro: è direzionata verso qualcosa, accumula tensione che chiede risoluzione, implica una meta che non è ancora stata raggiunta. Le misure centrali, con il prolungamento del V7 su quattro misure intere con il pedale al basso, costituiscono il momento di massima tensione dell'accomplishment armonico profondo: il terminus si avvicina, la struttura tonale lo indica con crescente urgenza, e l'ascoltatore che conosce il sistema tonale sente questa urgenza anche se la superficie ritmica continua a arpeggiare con la stessa omogenea continuità di prima. Le misure finali portano la risoluzione — la tonica in posizione fondamentale — che è il terminus dell'accomplishment armonico profondo.

Ciò che si ha, dunque, nel Preludio in Do è una struttura aspettuale a due livelli sovrapposti che devono essere tenuti analiticamente distinti: la superficie è atelica — ogni misura è fungibile, intercambiabile, omogenea rispetto alle altre, sospendibile senza incoerenza — mentre il livello armonico profondo è telico — c'è un terminus verso cui la struttura tonale converge con crescente urgenza man mano che ci si avvicina alla fine. La tensione cognitiva tra questi due livelli è precisamente ciò che produce l'effetto estetico peculiare del Preludio: la sensazione di una superficie assolutamente immobile e ripetitiva che eppure "va da qualche parte", che è sostenuta da una corrente profonda che la superficie non mostra ma che si sente scorrere sotto. Questo meccanismo di doppia stratificazione aspettuale — atelicità di superficie su telicità profonda — non è una peculiarità del Preludio bachiano ma un meccanismo pervasivo della musica tonale: lo stesso fenomeno si trova in quasi tutti gli sviluppi della forma-sonata e nei passaggi di variazioni su basso ostinato.

2.4 Gli accomplishments musicali: il processo verso il terminus e la sua gradazione

Gli accomplishments musicali sono eventi che si svolgono dinamicamente nel tempo verso un terminus intrinseco — eventi che "costruiscono verso" il loro punto di completamento in modo tale che l'interruzione prima del terminus produce un evento incompleto piuttosto che semplicemente un evento più breve. Il caso prototipico è la progressione cadenzale, e in particolare la cadenza autentica perfetta nel sistema tonale classico — la struttura che la teoria musicale da Rameau in poi ha riconosciuto come il gesto tonale per eccellenza, il dispositivo attraverso cui il sistema tonale afferma la sua gerarchia, proclama i suoi centri gravitazionali e definisce i confini delle strutture formali a tutti i livelli della scala.

Una cadenza autentica perfetta — nella sua forma più elementare ma non per questo meno eloquente, la progressione V7-I con la sensibile che risolve sulla tonica alla voce superiore e il basso che salta dalla dominante alla tonica in posizione fondamentale — è un accomplishment nel senso preciso e operazionalmente verificabile di Vendler. Ha fasi strutturalmente distinte che non sono fungibili: la presentazione dell'accordo di dominante con settima, la risoluzione sulla tonica. Ha una direzionalità interna inequivocabile: la sensibile "vuole" salire di semitono alla tonica, la settima dell'accordo di dominante "vuole" scendere di grado alla terza dell'accordo di tonica, il basso "vuole" saltare di quinta discendente alla tonica. Questi "voleri" non sono metafore romanticizzanti: sono la descrizione di proprietà cognitivamente reali del sistema tonale che ogni ascoltatore con esperienza della tonalità percepisce come forze direzionali reali, e che gli studi sperimentali di Bharucha e Stoeckig (1986) e di Bigand, Parncutt e Lerdahl (1996) hanno quantificato in termini di tempi di reazione e di aspettativa armonica. Ha un terminus intrinseco — la posizione fondamentale dell'accordo di tonica con la tonica alla voce superiore — che non è intercambiabile con nessun altro accordo senza distruggere l'identità categoriale dell'evento.

Un'interruzione prima del terminus — un arresto del V7 prima della risoluzione sulla tonica — produce una struttura percepita come incompleta nel modo preciso in cui è incompleto un cerchio non chiuso: non è una cadenza diversa, non è una cadenza "di tipo breve", non è una variante accorciata dello stesso evento. È un'aspettativa delusa, un processo aperto che non si è chiuso, una tensione che è stata generata e non è stata risolta. Questa percezione di incompletezza è indipendente dalla tradizione culturale dell'ascoltatore, nella misura in cui l'ascoltatore abbia familiarità con il sistema tonale: non si tratta di una convenzione stilistica che vuole che le cadenze si completino, ma di una proprietà strutturale del sistema tonale che il sistema stesso iscrive nell'evento.

Gli accomplishments musicali non si limitano alle cadenze: l'intera struttura della forma-sonata è strutturata come un accomplishment di scala superiore nel senso di Jackendoff. La ripresa non è semplicemente una terza sezione che ha la stessa struttura dell'esposizione: è il terminus intrinseco dell'intero processo formale, il punto in cui la dialettica tematica iniziata nell'esposizione raggiunge il proprio completamento nella tonalità principale. Un'interruzione della forma-sonata prima della ripresa — come succede in certi frammenti incompiuti di Schubert — produce una struttura incompleta che manca del suo terminus, non una struttura completa di tipo diverso.

La forma-sonata come accomplishment introduce però un problema che richiede qualificazione teorica: il problema dei livelli di scala nella struttura aspettuale musicale. Un brano può essere un accomplishment a livello macrostrutturale mentre le sue sezioni interne sono strutturate diversamente secondo la loro funzione.

2.5 Gli achievement musicali: l'istante che cambia tutto e reinterpreta il passato

Gli achievement musicali sono i momenti di cambiamento categoriale nella struttura sonora: gli istanti in cui un evento sonoro cambia tipo, in cui una qualità dell'ascolto si trasforma in un'altra qualità qualitativamente e categorialmente diversa, in cui lo stato cognitivo dell'ascoltatore subisce una discontinuità che non si svolge progressivamente ma avviene in un punto senza spessore temporale. Sono i più difficili da descrivere con precisione perché la loro caratteristica definitoria — l'istantaneità, la puntualità, l'assenza di fasi interne — si scontra con il fatto ovvio che la musica si svolge nel tempo e nessun evento sonoro è letteralmente privo di durata. Come si trasferisce al dominio musicale la categoria dell'achievement, che in linguistica descrive eventi come "arrivare", "trovare", "riconoscere" che sono definiti proprio dalla loro struttura puntuale?

La risposta — che riprende la soluzione proposta per il dominio linguistico da Moens e Steedman (1988) nel loro concetto di "event nucleus" — è che la puntualità degli achievement non è una proprietà fisica della durata dell'evento ma una proprietà strutturale: un achievement è un evento la cui struttura non ha parti rilevanti che appartengano alla sua definizione come evento. Non ha fasi che possano essere compiute in modo incompleto, non ha un processo preparatorio che faccia parte della struttura dell'achievement stesso piuttosto che di ciò che precede l'achievement. "Arrivare" non ha parti nel senso rilevante: non c'è una "prima fase dell'arrivare" distinta da una "seconda fase dell'arrivare". Il processo che porta all'arrivo — il viaggio — non appartiene alla struttura dell'achievement ma all'accomplishment che lo precede; l'achievement stesso è il punto terminale di quell'accomplishment, il momento della transizione di stato.

La risoluzione di una dissonanza sull'accordo consonante è un achievement nel senso più preciso: l'istante in cui la settima di dominante si risolve sulla terza della tonica non ha parti interne, non si svolge progressivamente — la dissonanza è dissonante in un momento, poi è consonante nel momento successivo, e tra i due momenti non c'è un intervallo di "parziale risoluzione" che sia strutturalmente parte dell'achievement. La realizzazione dell'inganno di una cadenza d'inganno — il momento in cui l'ascoltatore si accorge che invece della tonica è arrivato il sesto grado — è un achievement cognitivo di natura peculiare: non solo il sesto grado suona dove si aspettava la tonica, ma l'accordo di dominante che lo precede viene retroattivamente reinterpretato non come un accomplishment che ha raggiunto il suo terminus atteso ma come un accomplishment che ha deviato verso un terminus inatteso. L'achievement della cadenza d'inganno non è semplicemente l'apparizione del sesto grado: è il momento in

cui tutta la struttura armonica precedente viene reinterpretata retroattivamente, il momento in cui la storia armonica del brano fino a quel punto cambia senso.

Questo effetto retroattivo — il fatto che l'achievement non solo cambia lo stato presente ma ricolore il passato — è una delle proprietà più cognitivamente interessanti degli achievement musicali, e non ha un corrispondente diretto nella linguistica perché nel linguaggio gli achievement sono normalmente eventi che si proiettano verso il futuro, mentre nella musica certi achievement hanno un effetto precipuamente retroattivo. Gli studi sperimentali di Bharucha e Stoeckig (1986) sulla risposta cerebrale all'attesa armonica e la sua soddisfazione o violazione offrono una base neuroscientifica per la realtà cognitiva di questi achievement: i pattern di risposta corticale ai momenti di "risoluzione" — sia attesa che inattesa — mostrano una discontinuità che corrisponde alla struttura puntuale dell'achievement, non alla struttura graduale del processo.

2.6 L'aspetto nella musica vocale: il problema del doppio strato e il controaspetto compositivo

La musica vocale introduce nella teoria dell'aspetto musicale una complicazione teorica che non si presenta nella musica puramente strumentale: il fenomeno che si potrebbe chiamare il *doppio strato aspettuale*. Ogni brano vocale porta con sé, per il solo fatto di intonare un testo in linguaggio naturale, due livelli aspettuati sovrapposti e potenzialmente in conflitto: il livello dell'aspetto grammaticale del testo, determinato dalla morfologia e dalla sintassi del linguaggio naturale impiegato; e il livello dell'aspetto musicale della linea vocale e dell'armonia, determinato dalle scelte compositive. Quando questi due livelli sono sincronizzati — quando l'aspetto linguistico del testo e l'aspetto musicale della composizione sono della stessa categoria — si produce un effetto di rinforzamento aspettuale che aumenta la coesione cognitiva del brano. Quando sono dissincroni — quando l'aspetto linguistico punta in una direzione e l'aspetto musicale punta in un'altra — si producono tensioni cognitive peculiari che appartengono a uno spazio intermedio che questa sezione si propone di mappare.

Il caso paradigmatico di sincronizzazione aspettuale nella musica vocale è il recitativo secco del barocco italiano, e in particolare il recitativo operistico da Monteverdi a Scarlatti, da Handel a Vivaldi. Nel recitativo, la struttura frasale del testo determina la struttura aspettuale della musica con un'immediatezza quasi meccanica che la teoria aspettuale permette di descrivere con precisione strutturale. Un periodo sintattico che si conclude con un verbo al passato remoto — "corse", "raggiunse", "vinse" — è un periodo che descrive linguisticamente un evento telico: il passato remoto italiano, come il perfetto russo, presenta l'evento come compiuto, come raggiunto il suo terminus, come visto dall'esterno nella sua totalità conclusa. La musica del recitativo che intona questi verbi tende a cadenzare sulla stessa sillaba o immediatamente dopo: la cadenza autentica che segue il verbo telico è la manifestazione musicale della telicità linguistica, e la loro co-occorrenza non è illustrativa nel senso debole ma

strutturalmente omologa nel senso cognitivo preciso che entrambi — il verbo al passato remoto e la cadenza autentica — sono eventi telici che raggiungono un terminus. Al contrario, un verbo all'imperfetto — "camminava", "pensava", "desiderava" nel senso di processo in corso — è linguisticamente atelico, e la musica del recitativo lo intona spesso su armonie che non cadenzano ma si prolungano o modulano in modo non-telico.

Il caso più ricco e più interessante teoricamente si verifica però quando i due livelli aspettuali sono in conflitto intenzionale — quando il compositore sfrutta la possibilità di impostare un testo con una certa struttura aspettuale su una musica con una struttura aspettuale diversa, producendo quello che si può chiamare *controaspetto compositivo*: la scelta deliberata di contrastare l'aspetto del testo con l'aspetto della musica per ottenere effetti espressivi che nessuno dei due livelli da solo potrebbe produrre. Il Lied romantico tedesco — specialmente nel ciclo di composizioni da Schubert a Wolf che rappresenta il vertice della tradizione — è il dominio in cui il controaspetto compositivo è stato sviluppato con la maggiore coscienza tecnica e i risultati estetici più ricchi.

Nel *Winterreise* di Schubert, il quinto pezzo del ciclo — "Der Lindenbaum" — offre uno degli esempi più eloquenti di sincronizzazione aspettuale perfetta tra testo e musica: il testo descrive uno stato permanente — il tiglio sta lì, è una presenza immobile — e la musica risponde con la forma canzone nella tonalità di Mi maggiore che si ripete stroficamente senza modificazioni strutturali, una forma che ha struttura aspettuale di attività abituale nel senso di Comrie: ogni strofa ripete il pattern, come il tiglio ripete la sua presenza ad ogni passaggio del viandante. La sincronizzazione è perfetta: stativo linguistico e atelico-abituale musicale si sovrappongono producendo un effetto di staticità cognitiva che è molto più potente di quanto ciascuno dei due livelli produrrebbe da solo.

Il caso più cognitivamente interessante, e il più caratteristicamente schubertiano, è quello del pezzo "Im Dorfe" — dove il testo descrive i sogni dei dormienti con una qualità statico-atelica ("*Die Menschen schlafen auf ihren Kissen*") e la musica risponde con un ostinato di terzine al basso che è strutturalmente atelico (il pattern si ripete indefinitamente senza terminus intrinseco) ma che ha, nella sua ripetizione stessa, una qualità ossessiva e turbante che contrasta con la quiete stativa del testo. Qui il controaspetto è sottile: entrambi i livelli sono atelici, ma di qualità diversa — lo stativo del testo è quieto e immobile, l'atelico-iterativo della musica è inquieto e martellante — e questa differenza qualitativa all'interno dell'atelicità condivisa produce una dissonanza espressiva che è l'essenza della sensibilità schubertiana: la musica "sa" qualcosa che il testo non dice.

Il controaspetto più diretto e più evidente si trova invece in un celebre esempio mozartiano: l'aria "*Ach, ich liebte, war so glücklich*" (*Il ratto dal serraglio*, K. 384). Il testo al passato — "*ich liebte*", "*war so glücklich*" — è linguisticamente telico: il passato tedesco presenta l'azione come completamente conclusa, come vista dall'esterno nella sua

interezza. Ma Mozart intona questi verbi telici su una linea musicale volutamente sospesa e non-cadenzante: la linea vocale si protrae su armonie che non risolvono, che si prolungano in modo non-telico, che suggeriscono una qualità imperfettiva di processo in corso piuttosto che la qualità perfettiva di evento concluso. Il risultato percettivo è quello di una compiutezza linguisticamente dichiarata e musicalmente negata: il testo dice "ho amato e quella storia è finita" — evento telico concluso — la musica dice "sto ancora amando e quella storia non è veramente finita" — processo atelico in corso. Questa contraddizione tra il livello linguistico-aspettuale e il livello musicale-aspettuale è il meccanismo cognitivo attraverso cui Mozart esprime lo stato psicologico di Konstanze: sa intellettualmente che la storia è conclusa, lo dice con le parole, ma emotivamente non riesce a concluderla, e la musica rivela ciò che le parole nascondono.

Il madrigale rinascimentale italiano — da Marenzio a Gesualdo, dal Monteverdi madrigalistico a Luca Marenzio — pratica una sincronizzazione quasi sillabica tra aspetto grammaticale del testo e aspetto musicale: ogni parola significativa riceve il trattamento musicale che corrisponde alla sua categoria aspettuale, i verbi di azione telica cadenzano, i verbi di stato si prolungano, i verbi di moto generano figure scalari direzionate verso il loro terminus. Questa iperletteralità aspettuale è la caratteristica che rende i madrigali del Rinascimento italiano così ricchi di gesti illustrativi e al contempo così analizzabili dal punto di vista strutturale. Il corale luterano — nella sua forma più sistematica e più elaborata, quella bachiana — pratica invece una sincronizzazione che opera a livello di frase piuttosto che di parola: non è la morfologia verbale del singolo verbo che determina la struttura aspettuale della frase musicale ma la struttura sintattica del versetto nel suo arco retorico complessivo. Il canto gregoriano, con la sua indifferenza alla grammatica tonale funzionale e la sua preferenza per il moto modale per gradi congiunti che non genera tensione direzionale, tende a una quasi-neutralizzazione dell'aspetto musicale a livello di superficie, delegando interamente la struttura aspettuale al testo.

2.7 Obiezione principale: tutta la musica tonale è telica, poiché la tonica è sempre il terminus

L'obiezione più difficile, tecnicamente più impegnativa e analiticamente più produttiva alla proposta di una tassonomia aspettuale degli eventi musicali nel sistema tonale è quella che si formula più semplicemente così: nel sistema tonale, ogni progressione ha una tonica come punto di riferimento gravitazionale, e questa tonica funziona come terminus implicito di ogni evento armonico. In questo senso, la musica tonale sarebbe interamente telica: ogni evento punta verso la tonica, ogni tensione tende alla risoluzione, ogni dissonanza cerca la consonanza. La distinzione tra eventi telici e atelici proposta in questo capitolo sarebbe quindi o triviale (tutti gli eventi sono telici, e la differenza è solo di grado o di scala temporale) o arbitraria (si sceglie un terminus diverso dalla tonica come "definitorio" dell'evento specifico, ma questa scelta è soggettiva e non strutturalmente motivata).

Questa obiezione ha una versione analitica sofisticata nella teoria di Heinrich Schenker, nella sua formulazione più coerente: la tesi che tutta la musica tonale è strutturalmente telica al livello della sua struttura di fondo — la cosiddetta *Ursatz*, la struttura profonda che descrive il movimento dalla quinta o dalla terza del sistema alla tonica attraverso la progressione scalare discendente *Urlinie* e il movimento di basso I-V-I. In questa visione, ogni evento di superficie è una elaborazione — un dispiegamento nel tempo (*Auskomponierung*) — di questa struttura profonda, e la telicità dell'*Ursatz* permea tutta la musica tonale dall'inizio alla fine. Un ostinato di basso è telico quanto una cadenza autentica: è semplicemente un livello diverso dell'elaborazione dello stesso movimento profondo verso la tonica.

Questa risposta schenkeriana è internamente coerente e merita rispetto. Ma ha un limite che l'analisi aspettuale può mettere in luce con precisione: la teoria schenkeriana descrive la struttura del sistema tonale come sistema — la logica immanente del sistema, ciò che il sistema "è" nella sua idealità strutturale — non la struttura cognitiva degli eventi musicali come eventi percepiti in tempo reale da un ascoltatore in carne e ossa. L'*Ursatz* è una descrizione della logica immanente della tonalità, non di come l'ascoltatore processa la musica momento per momento. E la distinzione aspettuale che questa monografia propone non riguarda la struttura del sistema ma la struttura cognitiva dell'esperienza musicale.

La prova cognitiva più diretta che la distinzione aspettuale non è triviale viene dalla fenomenologia differenziale dell'ascolto. Ascoltare un ostinato di basso e ascoltare una progressione cadenzale sono esperienze del tempo musicale qualitativamente diverse che non si riducono a differenze di struttura armonica o di funzione tonale profonda. Un ascoltatore che sente un ostinato di basso non "aspetta la tonica" nel modo in cui un ascoltatore che sente un accordo di V7 aspetta la tonica: il suo stato cognitivo è diverso, il suo rapporto con il tempo musicale è diverso, la qualità della sua attesa è diversa. L'ascoltatore del V7 è in uno stato di tensione direzionale: sa che qualcosa sta per succedere, sa approssimativamente cosa, e la sua attenzione è orientata verso un punto futuro specifico. L'ascoltatore dell'ostinato di basso è in uno stato di immersione iterativa: non aspetta nessun punto futuro specifico perché l'ostinato potrebbe continuare indefinitamente. Queste due qualità di attenzione e di aspettativa sono cognitivamente diverse in modo verificabile: sono diverse nella fenomenologia soggettiva, diverse nei pattern di attivazione cerebrale che la neuroimaging musicale ha iniziato a mappare, diverse nelle risposte di sorpresa e di soddisfazione che producono quando le aspettative vengono soddisfatte o deluse.

La risposta definitiva all'obiezione è quindi la seguente: la distinzione aspettuale non contraddice la struttura telica profonda del sistema tonale, ma opera a un livello diverso di descrizione che è cognitivamente autonomo rispetto a quello schenkeriano. Il sistema tonale è telico a livello della struttura di fondo (la logica immanente del sistema tende sempre alla tonica come *terminus ultimum*); gli eventi musicali specifici hanno Aktionsart diverse a livello della struttura cognitiva di medio termine (la qualità dell'esperienza temporale immediata). Non c'è contraddizione tra questi due livelli,

così come non c'è contraddizione tra il fatto che una conversazione "tende" nel lungo periodo verso una conclusione e il fatto che alcune frasi della conversazione siano attività e altre siano accomplishments. La telicità di macro-livello coesiste con la variazione aspettuale di micro-livello, e sono precisamente i rapporti tra questi due livelli a produrre la ricchezza cognitiva dell'esperienza musicale.

2.8 Il confine tra le categorie: ambiguità aspettuale e oscillazione strutturale

Le quattro categorie aspettuali non formano un sistema di categorie mutuamente esclusive e prive di zone di confine. In linguistica, i casi di ambiguità aspettuale sono numerosi, ben documentati e analiticamente produttivi: "*she is living in Paris*" può essere il progressivo di un'attività (vivere a Parigi come processo in corso, con l'implicazione che non è permanente) o il progressivo di uno stato (essere residente a Parigi come condizione che si "sta" manifestando), e la scelta tra le due interpretazioni dipende dal contesto e dalle intenzioni comunicative. Analogamente, in musica ci sono eventi la cui categorizzazione aspettuale è genuinamente ambigua o dipende dal livello di analisi scelto, e queste ambiguità sono esse stesse analiticamente significative.

Il caso più interessante e più pervasivo è quello della variazione tematica, che oscilla strutturalmente tra l'accomplishment e l'attività a seconda della relazione che si stabilisce tra il tema e le sue variazioni. In un tema con variazioni del tipo classico — le variazioni delle sonate di Haydn, la Arietta dell'Op. 111 di Beethoven, le Variazioni Goldberg di Bach come struttura complessiva — ogni variazione è relazionata al tema come un'elaborazione del medesimo materiale: la variazione non "porta avanti" il processo verso un nuovo terminus ma ricrea il medesimo processo con qualità di superficie diverse. In questo senso, la struttura del tema con variazioni è atelica a livello di macro-forma: si può sempre aggiungere un'altra variazione senza che la struttura cambi tipo, e la fine delle variazioni è sempre una decisione compositiva esterna alla struttura dell'evento. Ma ogni singola variazione è, internamente, strutturata come un accomplishment se ha una propria cadenza finale che la chiude come unità autonoma, o come un'attività se le sue figure melodiche o armoniche non tendono verso un terminus locale. Il tema con variazioni è dunque un caso paradigmatico di oscillazione aspettuale: l'evento di macro-livello è atelico-iterativo, gli eventi di micro-livello sono (alcuni) telici, e la tensione tra questi due livelli produce la caratteristica qualità cognitiva di questa forma — la sensazione di immobilità dinamica, di moto nel posto, di cambiamento che non va da nessuna parte.

Un secondo caso di ambiguità strutturale è quello dell'improvvisazione nella tradizione della musica dotta occidentale prima dell'Ottocento — le fantasia, le toccata, i preludi non misurati, i Tombeau per liuto — che oscillano costantemente tra attività atelica (il gesto improvvisatorio che si svolge senza terminus predeterminato) e accomplishment telico (il momento in cui il gesto improvvisatorio converge su una cadenza e la raggiunge). La struttura aspettuale dell'improvvisazione libera è quella di un'alternanza irregolare tra zone di atelicità esplorativa e zone di telicità cadenzale, un'alternanza che

non segue nessuna grammatica formale predeterminata ma che risponde alle esigenze estemporanee dell'esecutore. Questa oscillazione — non la permanenza stabile in una categoria aspettuale ma il continuo passaggio tra categorie diverse — è la caratteristica cognitiva che distingue l'improvvisazione dalla composizione nel senso più profondo: la composizione può pianificare le strutture asettuali e la loro distribuzione nel tempo, l'improvvisazione le scopre nell'atto stesso del suonare.

2.9 I marcatori asettuali musicali: strutture che modificano la qualità temporale degli eventi

Fin qui si è parlato prevalentemente di Aktionsart musicale — delle proprietà asettuali intrinseche degli eventi musicali, quelle che appartengono alla loro struttura "lessicale" prima che qualsiasi dispositivo grammaticale intervenga a modificarle. Ma il framework aspettuale include anche la dimensione dell'aspetto grammaticale musicale, ovvero delle strutture musicali che modificano la qualità temporale di un evento in modo analogo ai morfemi asettuali delle lingue naturali.

La corona — il segno di prolungamento su una nota o un accordo — è forse il marcatore aspettuale musicale più evidente e più facilmente analizzabile. La sua funzione grammaticale aspettuale è quella di un imperfettivizzatore: trasforma un evento che nella sua struttura intrinseca è un momento di un processo in corso — il V7 come momento di tensione in un accomplishment cadenzale — in qualcosa di strutturalmente diverso: un oggetto temporale isolato, decontestualizzato dal suo continuum processuale, presentato come un momento che "vale" da solo, sospeso fuori dal tempo direzionale del processo che lo circonda. Un V7 con corona non è un V7 "più lungo" nel senso che la sua durata fisica aumenta: è un V7 che ha cambiato la sua qualità aspettuale, passando da essere un momento di un accomplishment direzionato verso la tonica a essere qualcosa di più simile a uno stato — la tensione prolungata, l'attesa sospesa, la dissonanza che non si muove verso la risoluzione ma rimane immobile, congelata.

La ripetizione letterale è un marcatore aspettuale di natura più complessa e ambivalente, che può operare in direzioni diverse a seconda del contesto strutturale. Nella sua forma più elementare — la doppia barra di ripetizione che prescrive di ripetere letteralmente una sezione intera — la ripetizione trasforma un accomplishment (la sezione con la sua cadenza finale) in un evento iterativo: la sezione si ripete come un pattern, e questa iterazione ha la struttura cognitiva di un'attività nel senso di Comrie, cioè di un processo che si reitera come proprietà strutturale dello spazio musicale. La prima volta che si suona la prima parte di un minuetto di Mozart, si esegue un accomplishment che si completa sulla cadenza finale; la seconda volta la stessa struttura, ormai nella sua iterazione, ha perso la qualità di unicità e ha acquisito la qualità di pattern.

Il ritardando finale — la progressiva decelerazione del tempo verso la cadenza conclusiva — è un marcatore aspettuale che "segnala" l'avvicinarsi del terminus di un

accomplishment: è il correlato musicale della prosodia discendente e della decelerazione articolatoria che nelle lingue parlate segnalano la fine dell'enunciato, la conclusione del turno conversazionale, l'arrivo al punto di chiusura di una struttura retorica. Non cambia la Aktionsart dell'evento — la cadenza finale era già un accomplishment telico prima del ritardando — ma la grammaticalizza in modo esplicito, la rende percettivamente saliente, la iscrive nella struttura sonora come elemento strutturale obbligatorio piuttosto che implicito. La sistematicità di questi marcatori aspettuuali in musica — la loro ricorrenza in contesti analoghi all'interno di una stessa tradizione, la loro funzione di modificare la qualità temporale degli eventi in modi prevedibili, la loro assenza in certe tradizioni e la loro presenza in altre — è la prova più robusta che l'aspetto musicale non è semplicemente una struttura cognitiva latente che l'analista scopre retrospettivamente, ma una struttura che le tradizioni musicali hanno sviluppato consapevolmente nel corso della loro storia.

3. La Tipologia Aspettuale delle Tradizioni Musicali

Dove viene grammaticalizzato il tempo interno?

3.1 La domanda tipologica e il suo statuto

Se la proposta del Capitolo II è corretta — se gli eventi musicali hanno Aktionsart nel senso di Vendler e se la musica dispone di marcatori aspettuuali che modificano la qualità temporale degli eventi in modo analogo ai morfemi grammaticali delle lingue naturali — si apre immediatamente una domanda tipologica che ha la stessa struttura della domanda posta da Comrie (1976) nei confronti delle lingue naturali: le diverse tradizioni musicali grammaticalizzano l'aspetto in modi sistematicamente diversi, privilegiando certi tipi di Aktionsart come nuclei strutturali e relegando altri alla periferia? Esistono, in altri termini, tradizioni musicali che hanno una preferenza sistematica per l'atelicità, così come esistono lingue che non grammaticalizzano l'aspetto morfologicamente e che quindi lasciano la struttura temporale degli eventi determinata esclusivamente dal contesto? E al contrario, esistono tradizioni che grammaticalizzano la telicità in modo così pervasivo e obbligatorio da rendere ogni evento armonico strutturalmente orientato verso un terminus specifico?

La risposta che questo capitolo propone è affermativa su entrambi i fronti, e si articola attraverso l'analisi di sei tradizioni musicali scelte per la loro distanza reciproca e per la chiarezza con cui manifestano preferenze aspettuuali diverse: il contrappunto rinascimentale e barocco, il classicismo viennese, il romanticismo tardo, il minimalismo americano, il jazz nelle sue varianti bebop e modale, e infine la musica spettrale ed

elettroacustica, che grammaticalizza l'aspetto attraverso la dimensione timbrica anziché quella armonica. Le tradizioni non-occidentali — che richiedono un trattamento più lungo e più cauto — sono riservate al Capitolo V.

Una precisazione metodologica è necessaria prima di procedere. Quando si dice che una tradizione "grammaticalizza" un tipo di aspetto, si intende che quella tradizione ha sviluppato strutture obbligatorie — non opzionali, non puramente stilistiche — che determinano la qualità temporale degli eventi musicali in modo sistematico. La grammaticalizzazione è più forte della semplice preferenza stilistica: è l'istituzionalizzazione strutturale di una scelta aspettuale come norma del sistema, esattamente come la grammaticalizzazione del perfettivo/imperfettivo in russo non è una preferenza stilistica dei parlanti russi ma una caratteristica obbligatoria della loro morfologia verbale. Questa distinzione — tra grammaticalizzazione obbligatoria e preferenza stilistica — è la chiave per distinguere le differenze tipologiche dalle differenze stilistiche superficiali.

3.2 Il contrappunto rinascimentale: l'aspetto come gesto, non come struttura

Il contrappunto modale della tradizione palestriniana si rivela, nella prospettiva aspettuale, come un sistema che non grammaticalizza l'aspetto in modo sistematico ma lo lascia emergere dal gesto locale. Questa è un'affermazione che richiede elaborazione precisa, perché potrebbe sembrare contraddittoria con la centralità delle clausole cadenzali nel sistema rinascimentale — clausole che, come si è visto nel Capitolo II, hanno la struttura di accomplishments telici. Come può un sistema basato su clausole teliche essere un sistema che "non grammaticalizza" l'aspetto?

La risposta sta nel grado di obbligatorietà e di pervasività delle strutture teliche nel sistema. Nel contrappunto palestriniano, le clausole cadenzali sono strutture locali che articolano il percorso modale — i punti di arrivo su *finalis*, *confinalis* e terza del modo — ma tra una clausola e l'altra il tessuto contrappuntistico è prevalentemente atelico: il moto per gradi congiunti, l'imitazione libera, le note di passaggio e le sincopi preparate producono un flusso sonoro che non è orientato verso nessun terminus locale specifico ma "scorre" tra un punto di articolazione e il successivo. Il contrappunto palestriniano alterna momenti di telicità (le clausole) e zone di atelicità (i percorsi interstitiali), e non esiste una grammatica che determini in modo obbligatorio la qualità aspettuale del percorso tra una clausola e l'altra.

Questo regime aspettuale — telicità puntuale in luoghi strutturalmente determinati, atelicità diffusa nel tessuto connettivo — è analogo, nella tipologia di Comrie, alle lingue che grammaticalizzano l'aspetto solo in certi contesti grammaticali specifici lasciando il resto del tessuto in uno stato aspettualmente non marcato. Il contrappunto rinascimentale è un sistema a bassa densità di grammaticalizzazione aspettuale: l'aspetto è presente ma non è il principio organizzativo dominante del sistema.

L'apertura del mottetto *Sicut cervus* di Palestrina (ca. 1584) esemplifica con precisione la struttura aspettuale di questo sistema. Il soggetto melodico — un'ascesa di quarta seguita da un ornamento discendente — ha internamente la struttura di un accomplishment: ascende verso un terminus, e questo movimento ha direzionalità. Ma la sua imitazione nelle quattro voci sfasate produce un tessuto complessivo che è atelico: le quattro traiettorie si sovrappongono in modo tale che in ogni momento c'è sempre almeno una voce nel mezzo del suo percorso ascendente, un'altra in discesa, un'altra che inizia il soggetto. Non c'è un momento in cui tutte le voci raggiungono simultaneamente il loro terminus: il tessuto imitativo crea una perpetua atelicità di macro-livello anche se ogni singola voce è localmente telica. Questo è il meccanismo aspettuale fondamentale del contrappunto imitativo rinascimentale: la sovrapposizione di eventi localmente telici in un tessuto complessivamente atelico.

Il contrappunto bachiano introduce una novità fondamentale rispetto a questa struttura: la subordinazione obbligatoria di tutte le strutture aspetuali locali alla struttura tonale di macro-livello, che è sempre un accomplishment telico orientato verso la cadenza. In Bach, non esiste un tessuto "aspettualmente non marcato" tra una cadenza e l'altra: ogni accordo è funzionale in modo che ne determina il contributo alla direzionalità dell'accomplishment complessivo, e questa funzionalità è strutturalmente obbligatoria, non opzionale. Il contrappunto bachiano ha dunque una densità di grammaticalizzazione aspettuale molto superiore a quello palestriniano: l'aspetto telico è presente non solo nei momenti cadenzali ma in ogni momento della struttura armonica, che è costantemente orientata verso il terminus della cadenza successiva. Questo è il primo gradino verso il sistema classico che ha sviluppato la grammatica della telicità al suo massimo grado.

3.3 Il classicismo viennese: la grammatica della telicità come sistema

Il classicismo viennese di Haydn, Mozart e del primo Beethoven è, nell'ipotesi proposta qui, il sistema musicale che ha grammaticalizzato la telicità nel modo più sistematico, obbligatorio e raffinato che la storia della musica occidentale conosca. Non si sta dicendo che i compositori classici abbiano "scelto" di essere telici nel senso in cui un parlante sceglie di usare il perfettivo. Si sta dicendo che il sistema armonico-formale del classicismo viennese ha la struttura di un sistema di grammaticalizzazione aspettuale, nel quale ogni elemento strutturale — l'accordo, la frase, il periodo, la sezione, la forma — ha una funzione aspettuale determinata dal suo posto nella gerarchia formale, e questa funzione è obbligatoria nel senso che deviare da essa produce effetti marcati, sorprese o violazioni che l'ascoltatore percepisce come infrazioni di una norma.

Il meccanismo fondamentale di questa grammaticalizzazione è la gerarchia cadenzale — il sistema di cadenze di diverso peso strutturale che articola la forma a livelli diversi di scala. La teoria della forma musicale classica, da Koch (1793) a Ratz (1951) a Caplin (1998), ha descritto questa gerarchia in termini puramente formali. Ma nella

prospettiva aspettuale, questa gerarchia è una grammatica di telicità graduata: ogni livello della gerarchia corrisponde a un grado diverso di "forza del terminus", a un grado diverso di completamento dell'evento. Una cadenza autentica perfetta (PAC) alla fine di un'esposizione sonatistica porta telicità massimale per quella sezione; una cadenza di semiposa (HC) al termine di una transizione porta telicità minima, lasciando aperta la direzionalità verso ciò che segue; una cadenza d'inganno (DC) nega il terminus atteso e reindirizza il processo verso un terminus alternativo. Questa scalatura non è libera: è determinata dalla posizione formale, il che significa che l'aspetto è grammaticalizzato nel senso preciso che la qualità aspettuale degli eventi è obbligatoriamente determinata dalla struttura del sistema.

Il teorico che ha descritto più accuratamente questa grammatica senza usare il termine "aspetto" è William Caplin, nel suo *Classical Form* (1998). Caplin distingue tra funzioni formali "tight-knit" e "loose" — funzioni che "stringono" (producono eventi fortemente telici) e funzioni che "allentano" (producono eventi debolmente telici o atelici) — e mostra come queste funzioni siano distribuite sistematicamente nella forma sonata. In termini aspettuali, la frase classica di Caplin è strutturata come una sequenza stato-attività-accomplishment: la presentazione (le prime quattro misure) è prevalentemente stativa-atelica, l'armonia si muove poco, il materiale tematico è esposto senza direzionalità verso un terminus specifico; la continuazione (le misure successive) è dinamico-atelica, il ritmo aumenta, l'armonia si muove più rapidamente, ma ancora senza tendere verso un terminus specifico; la funzione cadenzale (le misure conclusive) è telico-accomplishment, la progressione ii6-V7-I costituisce l'evento cadenzale che completa la frase. Questa sequenza — da stativo ad atelico a telico — è la struttura aspettuale fondamentale della frase classica, e si ripete, con variazioni, a tutti i livelli della gerarchia formale.

L'analisi della Sonata in do maggiore K. 545 di Mozart chiarisce questa proposta con una chiarezza quasi didattica. Le prime quattro misure — presentazione nella terminologia di Caplin — sono costruite su un'armonia di tonica in do maggiore con figurazione di basso albertino e una melodia scalare alla voce superiore. Aspettualmente, queste quattro misure sono statiche-ateliche: l'armonia non si muove, la melodia scende per gradi senza tendere verso nessun terminus specifico. Le misure cinque-otto introducono un movimento armonico più rapido con cromatismi: aspettualmente sono atelico-dinamiche, c'è movimento e direzione ma ancora non si è in un evento la cui struttura tenda verso un terminus specifico. Le misure successive segnano il passaggio da attività ad accomplishment in corso: la figurazione di semiminime accelera, l'armonia attraversa iv-ii-V7, e si sente chiaramente che ci si sta avvicinando a qualcosa. Le misure conclusive completano l'accomplishment: la progressione ii6-V7-I si compie con la terza nel soprano alla tonica finale. Questa sequenza stato-attività-accomplishment-achievement è la struttura aspettuale della frase classica, e nella K. 545 di Mozart è realizzata con una trasparenza che la rende un caso paradigmatico.

3.4 Il romanticismo tardo: l'aspetto sospeso tra telicità e cromatismo

Il romanticismo tardo — Wagner, Brahms, Bruckner, il primo Schoenberg — introduce nel sistema classico una complicazione aspettuale che è al tempo stesso il suo sviluppo più raffinato e il germe della sua dissoluzione. La grammaticalizzazione della telicità, che nel classicismo viennese era chiara e gerarchicamente ordinata, comincia a incepparsi in due modi distinti ma complementari. Il primo modo è la sospensione indefinita del terminus attraverso il cromatismo che ritarda sistematicamente la risoluzione: l'intera struttura del *Tristan* wagneriano è, da questo punto di vista, un unico accomplishment di scala enorme il cui terminus viene procrastinato per quattro ore attraverso una successione di risoluzioni negate, ricontestualizzazioni armoniche e deviazioni che non negano la struttura telica di fondo ma la distendono fino al limite della sopportabilità. Il secondo modo è la coesistenza di traiettorie armoniche multiple che hanno terminus diversi e incompatibili, producendo una forma di diatesi aspettuale ambigua che anticipa la polifonia aspettuale del Capitolo IV.

Nel sistema brahmsiano, questa tensione si manifesta in una forma peculiare che si potrebbe chiamare telicità contrastata: Brahms costruisce accomplishments di grande scala con terminus chiaramente impliciti ma poi li ritarda, li circoscrive, li maschera attraverso una complessità metrico-ritmica che produce l'effetto di una direzionalità sempre presente ma mai pienamente visibile in superficie. L'ascoltatore di Brahms sente costantemente che il processo è orientato verso qualcosa — la grammaticalizzazione della telicità classica è ancora attiva come norma di fondo — ma la superficie è abbastanza densa da oscurare il momento esatto in cui il terminus sarà raggiunto. Questa opacità del terminus non è atelicità nel senso del minimalismo: è telicità ad alta densità di resistenza, un accomplishment il cui cammino verso il terminus è pieno di ostacoli strutturali che rendono il completamento tanto più soddisfacente quanto più è stato ostacolato.

3.5 Il minimalismo americano: la grammatica dell'atelicità

Se il classicismo viennese è il sistema che ha portato la grammaticalizzazione della telicità al suo massimo grado di raffinatezza e obbligatorietà, il minimalismo americano — nelle sue forme più coerenti e sistematiche, da *In C* di Terry Riley (1964) alle opere di Philip Glass degli anni Settanta e all'early Steve Reich — è il sistema che ha grammaticalizzato sistematicamente l'atelicità, costruendo un'estetica e una tecnica compositiva fondate sulla rimozione strutturale del terminus come categoria organizzativa del discorso musicale.

In C di Riley è il manifesto del minimalismo aspettuale nel senso più preciso. Il brano consiste di cinquantatré brevi moduli melodici che i suonatori eseguono in sequenza, ognuno a proprio tempo, ripetendo ogni modulo il numero di volte che ritiene opportuno prima di passare al successivo. Non esiste un momento in cui l'esecuzione

si "completa" nel senso in cui si completa la frase classica. Ogni suonatore decide autonomamente quando passare al modulo successivo, e l'insieme termina quando tutti i suonatori hanno raggiunto l'ultimo modulo — ma anche questo è una decisione pratica, non una chiusura strutturale. Se si aggiungessero ulteriori ripetizioni dell'ultimo modulo, l'evento non sarebbe più "completo" di quanto fosse già: sarebbe semplicemente più lungo. Questa è la definizione operativa dell'atelicità.

La grammaticalizzazione dell'atelicità in *In C* non è solo nella struttura formale del brano ma nei meccanismi compositivi che costruiscono i singoli moduli. Ogni modulo è una breve figura melodica o ritmica costruita in modo da essere ripetibile senza creare aspettativa di terminus: le figure si chiudono su note che permettono la ripetizione diretta senza implicazione direzionale forte. Questo meccanismo di re-atelicizzazione per iterazione obbligatoria è uno dei contributi compositivi più originali del minimalismo, e ha un preciso analogo nella grammatica aspettuale delle lingue naturali: nelle lingue slave, il perfettivo di un verbo può essere "imperfettivizzato secondariamente" mediante un processo morfologico che trasforma l'evento telico in un'attività abituale. Il minimalismo opera una "imperfettivizzazione secondaria" attraverso l'iterazione: prende figure che hanno struttura telica locale e le re-atelicizza attraverso la ripetizione obbligatoria, trasformando ogni terminus locale in un nodo di un processo abituale-iterativo.

Glass, nelle sue opere degli anni Settanta, porta questa grammatica dell'atelicità a una scala architettonica. Le sezioni di *Einstein on the Beach* si estendono per decine di minuti su armonie che si muovono lentamente, senza la direzionalità cadenzale del sistema tonale classico. La struttura aspettuale di queste sezioni è quella di un'attività di scala superiore: c'è movimento armonico, ma questo movimento non ha terminus intrinseco, non tende verso una cadenza specifica che farebbe parte della struttura dell'evento. Le sezioni di *Einstein* finiscono perché il libretto lo richiede, non perché la struttura musicale abbia raggiunto il suo terminus. Questa è grammaticalizzazione dell'atelicità a livello macrostrutturale.

Reich, in *Music for 18 Musicians* (1976), introduce una variante della grammatica atelica che ha una struttura aspettuale più complessa. I cicli di pulsazione sovrapposti producono pattern emergenti che cambiano gradualmente nel tempo: c'è una teleologia nella struttura del brano — si va da uno stato iniziale a uno stato finale attraverso trasformazioni graduali — ma questa teleologia non è costruita attraverso la direzionalità cadenzale. Essa produce una struttura aspettuale gerarchicamente invertita rispetto al classicismo viennese: nel classicismo, gli eventi locali sono telici (le frasi cadenzano) e la forma globale è costruita dall'accumulo di telicità locali; in Reich, gli eventi locali sono atelici (i cicli si ripetono senza terminus) e la forma globale ha una telicità di macro-livello prodotta dalla trasformazione graduale. Questo rovesciamento della gerarchia aspettuale è una delle innovazioni strutturali più profonde del minimalismo.

3.6 Il jazz bebop: la grammatica dell'accomplishment rapido

Il jazz bebop — nelle sue forme prototipiche degli anni Quaranta con Charlie Parker, Dizzy Gillespie e Thelonious Monk — costituisce un terzo regime aspettuale che si può caratterizzare come la grammatica dell'accomplishment rapido: un sistema in cui la densità di cadenze, la velocità delle modulazioni e la ricchezza armonica producono una successione frenetica di eventi telici a piccola scala, con pochi momenti di vera atelicità e nessuna struttura stativa di durata significativa.

Nel bebop, la progressione di trentadue misure del "rhythm changes" è la struttura di riferimento, e la sua analisi aspettuale rivela la grammatica del sistema. Il chorus di trentadue misure è un accomplishment di macro-livello con inizio, svolgimento e terminus. Ma all'interno di questo accomplishment di macro-livello, ogni quattro misure contengono almeno una cadenza locale — spesso una progressione ii7-V7-I che completa un accomplishment di micro-livello — e tra una cadenza di micro-livello e la successiva non c'è quasi spazio per l'atelicità: il ritmo armonico è così rapido che ogni accordo è già parte della struttura direzionale dell'accomplishment successivo.

Questa densità di telicità a piccola scala ha conseguenze dirette per la pratica dell'improvvisazione. Un improvvisatore bebop non può permettersi periodi di atelicità estesa — di figurazioni non direzionali che "orbitano" senza tendere verso un terminus specifico — perché la progressione di accordi cambia troppo rapidamente e ogni accordo ha già una funzione determinata nell'accomplishment che sta costruendo. L'improvvisatore deve essere costantemente "in transito" verso il terminus del prossimo accordo: ogni nota della sua improvvisazione è parte di un accomplishment locale che si completa sull'accordo successivo. La frase che i teorici del jazz chiamano "bebop line" — una scala rapida di sedicesimi che percorre un ampio registro, attraversa accordi multipli e si conclude su una nota di target sull'accordo di risoluzione — è aspettualmente un accomplishment di grande intensità: il suo terminus è il punto verso cui tutta la linea è strutturalmente orientata.

Il jazz modale — già in nuce nell'album *Kind of Blue* di Miles Davis (1959), e portato alle sue conseguenze più estreme dall'Impulse! period di John Coltrane — ha un regime aspettuale che si colloca tra il minimalismo e il bebop in modo strutturalmente illuminante. Non è puramente atelico come Glass, non è densamente telico come Parker, ma usa l'aspetto in modo che si potrebbe chiamare "sospeso" — un regime in cui la qualità aspettuale degli eventi è deliberatamente mantenuta in uno stato di indeterminazione tra telicità e atelicità. Le improvvisazioni di Coltrane del periodo 1961-1965 iniziano come attività (senza chiara direzionalità verso un terminus) e poi sviluppano una direzionalità locale che si completa o si dissolve prima di raggiungere un terminus definitivo. Il terminus locale — il punto a cui la frase sembrava tendere — viene spesso eluso o trasformato in un nuovo punto di partenza, producendo quella qualità di sospensione perpetua che è la caratteristica più immediatamente percepibile dell'improvvisazione coltraniana matura. Questo regime di aspetto sospeso — il rifiuto sistematico di completare gli accomplishments che si iniziano — è una forma di grammaticalizzazione dell'ambiguità aspettuale che non ha un preciso analogo nelle tipologie linguistiche di Comrie.

3.7 La musica spettrale ed elettroacustica: l'aspetto nel timbro

La musica spettrale — sviluppata a Parigi intorno all'IRCAM negli anni Settanta e Ottanta da compositori come Gérard Grisey, Tristan Murail e Hugues Dufourt — introduce nella teoria dell'aspetto musicale una dimensione che nessuna delle tradizioni precedentemente analizzate aveva esplorato sistematicamente: la grammaticalizzazione dell'aspetto attraverso la dimensione timbrica piuttosto che attraverso quella armonica. In questo sistema, la distinzione telico/atelico non è realizzata attraverso la direzione delle progressioni armoniche verso una tonica ma attraverso la struttura temporale del suono come oggetto fisico-acustico: l'attacco di un suono, il suo svolgimento nel tempo, la sua decadenza verso il silenzio.

In *Partiels* di Grisey (1975), il pezzo che più compiutamente illustra il programma spettralista, la struttura formale è costruita interamente su processi di trasformazione timbrica che hanno la struttura aspettuale degli accomplishments: un suono inizia con una qualità timbrica data — un suono di trombone nella sua complessità spettrale completa — e si trasforma progressivamente, attraverso l'orchestrazione degli armonici naturali del suono originario, verso stati timbrici sempre più rarefatti o sempre più densi. Questa trasformazione ha un terminus intrinseco — lo stato di massima rarefazione o di massima densità che la trasformazione "cerca" nel corso del suo svolgimento — e una direzionalità interna che l'ascoltatore percepisce come orientazione verso qualcosa di non ancora raggiunto. Il terminus del processo spettrale non è una tonica nel senso armonico-funzionale ma uno stato di purezza timbrica (un suono sinusoidale puro, il limite astratto della serie armonica) o di densità massima (il rumore bianco, l'altro limite). La struttura aspettuale è preservata nella sua essenza — processo direzionato verso terminus intrinseco — ma è realizzata attraverso un dominio radicalmente diverso.

La musica elettroacustica concreta — da Pierre Schaeffer agli *études* di Stockhausen, dalla tradizione di musica su nastro allo studio del suono digitale — presenta ancora un diverso regime aspettuale. Poiché i materiali sonori dell'arte concreta sono estratti dalla realtà acustica quotidiana e poi manipolati, essi portano con sé la struttura aspettuale degli eventi fisici che li hanno generati: il suono di una porta che si chiude è un achievement (puntuale, cambiamento di stato irreversibile); il suono di una fontana è un'attività (omogenea, senza terminus); il suono di un motore che si avvia è un accomplishment (processo direzionato verso il regime a pieno regime). La composizione elettroacustica, in questa prospettiva, è anche una composizione di strutture aspetuali estratte dal mondo fisico e riorganizzate secondo logiche che possono conservare, destabilizzare o trasformare le loro Aktionsart originarie. La coercion aspettuale — il meccanismo per cui il contesto forza un evento in una Aktionsart diversa dalla sua — è particolarmente evidente nella musica concreta: un suono di porta che si chiude, ripetuto venti volte in successione rapida, perde la sua struttura di achievement e acquisisce la struttura di un'attività iterativa, esattamente come un accomplishment classico ripetuto acquisisce la struttura di un'attività abituale.

3.8 Obiezione: le differenze tipologiche sono differenze di stile, non di grammatica

L'obiezione più sostanziale alla proposta tipologica sviluppata in questo capitolo è quella che si potrebbe formulare come segue: le differenze tra il classicismo viennese, il minimalismo e il bebop nella distribuzione della telicità e dell'atelicità non sono differenze di grammatica aspettuale obbligatoria ma differenze di stile compositivo, di scelta estetica, di preferenza idiomatica. Il fatto che Mozart cadenzasse frequentemente e Glass non cadenzasse quasi mai non implica che i due sistemi abbiano grammatiche aspettuali diverse: implica semplicemente che Mozart aveva una preferenza stilistica per la cadenza e Glass non l'aveva.

Questa obiezione ha una prima forza nell'evidenza della variazione individuale: anche all'interno del classicismo viennese, si trovano compositori che cadenzano più o meno frequentemente, che usano le cadenze sospese e d'inganno con diversa libertà. Se la struttura aspettuale del classicismo fosse davvero grammaticalizzata nel senso dell'obbligatorietà, ci si aspetterebbe meno variazione individuale. Questa variazione suggerisce che si tratta di preferenze stilistiche piuttosto che di norme grammaticali.

La seconda forza dell'obiezione è la continuità storica: le tradizioni musicali non si sostituiscono bruscamente l'una all'altra come le lingue si distinguono nettamente tra loro, ma si trasformano gradualmente, con zone di transizione in cui i regimi aspettuali si mescolano. Il passaggio dal cromatismo romantico al post-tonale, o dal bebop al jazz modale, non sono passaggi da una grammatica aspettuale a un'altra nel modo in cui il russo e l'italiano hanno morfologie aspettuali distinte: sono transizioni graduali in cui i vecchi meccanismi coesistono con quelli nuovi per periodi lunghi.

Queste obiezioni sono giuste nell'identificare i limiti della tipologia ma traggono conclusioni troppo forti. La prima confonde la variazione nei parametri di un sistema con la variazione del sistema stesso: all'interno del classicismo viennese, la struttura gerarchica della cadenza — la grammatica della telicità graduata — è sistematicamente presente in tutti i compositori del periodo, nonostante le variazioni individuali nella frequenza e nella qualità delle cadenze. Un'analisi sistematica delle sonate per pianoforte di Haydn, Mozart e Beethoven troverebbe che la distinzione PAC/IAC/HC è universalmente osservata nelle posizioni strutturali che la grammatica assegna loro — la fine di una sezione formale principale è sempre una PAC, la fine di una transizione è sempre una HC o IAC — con una sistematicità che non è spiegabile come preferenza individuale ma solo come norma del sistema. La seconda obiezione confonde la gradualità delle transizioni storiche con la continuità del sistema: il fatto che le tradizioni si trasformino gradualmente non implica che le distinzioni tipologiche tra di esse siano arbitrarie.

La risposta definitiva è dunque: la grammaticalizzazione aspettuale nelle tradizioni musicali non è assoluta e binaria come la distinzione perfettivo/imperfettivo del russo, ma è reale e sistematica nel senso che le norme strutturali di ciascun sistema

determinano in modo obbligatorio la qualità aspettuale degli eventi in certi contesti strutturali chiave. Un brano che viola sistematicamente queste norme — che porta una PAC nel mezzo di una transizione, o che interrompe un processo minimalista con una cadenza autentica netta — viene percepito dall'ascoltatore addestrato come marcato, anomalo, stilisticamente incoerente. Questa percepibilità della violazione è la prova che la grammaticalizzazione è reale: dove c'è norma violabile, c'è grammatica.

4. La Diatesi Aspettuale nella Polifonia

Dissincronia, coercion e il tempo plurale del contrappunto

4.1 Il problema della simultaneità aspettuale

Tutto ciò che si è detto finora sull'aspetto musicale si è svolto implicitamente su una singola linea temporale: un evento armonico ha una certa Aktionsart, un certo grado di telicità, un certo tipo di marcatura aspettuale, e queste proprietà si dispiegano nel tempo come proprietà di un evento singolo. Anche le analisi della polifonia nel Capitolo III hanno considerato la struttura aspettuale del brano come un tutt'uno — il tessuto imitativo complessivo di Palestrina, la struttura gerarchica cadenzale di Mozart, i cicli sovrapposti di Reich — senza affrontare la questione più difficile e più originale che la polifonia pone alla teoria dell'aspetto: che cosa succede quando voci diverse, che si svolgono simultaneamente nello stesso spazio temporale, sono in stati aspettuuali diversi?

Il problema può essere formulato con precisione senza ricorrere a formalismi: in un contrappunto a quattro voci, è possibile — e in realtà frequentissimo — che in un dato momento la voce di soprano sia nel mezzo di un accomplishment (sta costruendo verso il terminus della sua cadenza locale), il contralto sia in uno stato (tiene una nota lunga su un pedale armonico), il tenore sia in un'attività (percorre una sequenza atelica), e il basso sia appena arrivato all'achievement della risoluzione di una dissonanza. Quattro voci, quattro Aktionsart diverse, lo stesso istante temporale. Questa coesistenza di stati aspettuuali diversi nello stesso spazio temporale — che chiamo *diatesi aspettuale*, prendendo in prestito il termine dalla grammatica per indicare la relazione tra una voce e il tipo di evento in cui si trova — è la caratteristica più originale e cognitivamente più ricca della polifonia, e nessuna teoria musicale disponibile ha gli strumenti per descriverla con precisione.

Il termine "diatesi aspettuale" richiede giustificazione. In grammatica, la diatesi descrive la prospettiva da cui un evento è presentato — il soggetto è agente (diatesi attiva), paziente (diatesi passiva), beneficiario (diatesi media) — e questa prospettiva

modifica la struttura dell'evento senza cambiarne il contenuto proposizionale. Nella polifonia, la diatesi aspettuale descrive il tipo di evento in cui ciascuna voce si trova — accomplishment, attività, stato, achievement — e questo tipo di evento è la "prospettiva temporale" dalla quale ciascuna voce partecipa al tessuto complessivo. Come in grammatica la diatesi attiva e passiva coesistono in costruzioni perifrastiche complesse, nella polifonia le diatesi aspetuali di voci diverse coesistono nel tempo, interagendo e producendo tensioni che non sono riducibili né alla tensione armonica né alla tensione ritmica ma appartengono a una categoria propria.

Che questa categoria sia cognitivamente reale — che la dissincronia aspettuale produca una tensione percepibile distinta da altri tipi di tensione — è la tesi centrale di questo capitolo, argomentata attraverso tre approcci convergenti: l'analisi fenomenologica dell'ascolto, l'analisi strutturale di brani specifici, e il collegamento con la teoria neuroscientifica del tempo musicale.

4.2 La dissincronia aspettuale come fonte di tensione: fenomenologia

Quando si ascolta un contrappunto in cui le voci si trovano in stati aspetuali diversi, si sperimenta una qualità del tempo musicale che non è semplicemente la somma delle qualità temporali delle singole voci ma qualcosa di strutturalmente diverso. Consideriamo un caso semplice: un basso che tiene un pedale di tonica (stato) mentre le voci superiori costruiscono una progressione cadenzale (accomplishment in corso). Fenomenologicamente, si sperimenta un'asimmetria temporale: il basso "è" in un tempo stabile, fermo, che non progredisce verso nessun terminus; le voci superiori "stanno andando" verso un terminus, il loro tempo è direzionale e orientato. Questa asimmetria non è semplicemente la differenza tra una voce lenta e una voce rapida: è una differenza nella qualità del tempo. Il basso è in un "presente esteso" (lo stato atelico che "vale" senza tendere), le voci superiori sono in un "futuro anticipato" (il processo telico che "costruisce verso"). La tensione tra questi due regimi temporali — la coesistenza di un "presente" e di un "futuro" nella stessa struttura polifonica — è la tensione della diatesi aspettuale.

Questo tipo di tensione è distinto dalla tensione armonica nel senso lerdahliano, che misura la distanza dalla tonica nello spazio tonale, perché non dipende dalla qualità degli accordi ma dalla struttura temporale degli eventi. Un pedale di tonica con un accomplishment cadenzale sopra di esso ha bassa tensione armonica (tutti gli accordi sono funzionalmente vicini alla tonica) ma alta tensione aspettuale (la coesistenza di stato e accomplishment crea un'asimmetria temporale che richiede risoluzione). Al contrario, un'armonia cromaticamente distante dalla tonica ma strutturata come puro stato ha alta tensione armonica ma bassa tensione aspettuale. L'irriducibilità delle due tensioni è la prova che sono di tipo diverso.

L'irriducibilità è dimostrata anche rispetto alla tensione ritmica: un pedale di tonica con un accomplishment cadenzale sopra può avere struttura ritmica perfettamente

regolare e simmetrica, senza nessuna tensione ritmica, e tuttavia produce la tensione aspettuale della dissincronia. La tensione aspettuale è una categoria propria, non derivabile né dalla teoria armonica né dalla teoria ritmica: è la tensione del tempo plurale, del fatto che le diverse voci abitano tipi di tempo qualitativamente diversi e incompatibili.

4.3 La fuga in do diesis minore del WTC I: analisi sistematica della diatesi aspettuale

La fuga in do diesis minore dal primo volume del *Clavicembalo Ben Temperato* (BWV 873) è scelta per questa analisi perché è un brano di media complessità — tre voci, quarantadue misure — che manifesta una varietà eccezionalmente ricca di configurazioni di diatesi aspettuale. Il soggetto della fuga è costruito su un'ascesa cromatica di quarta, seguita da una figurazione di arpeggio discendente e da una nota tenuta sulla dominante. Aspettualmente, il soggetto è un accomplishment di media scala: ha una fase ascendente direzionalmente orientata verso il terminus cromatico, una fase ornamentale, e una fase di "arrivo" sulla dominante che prepara la risposta.

La risposta — l'entrata della seconda voce sulla dominante — è lo stesso accomplishment trasposto, e la controesposizione produce il primo momento di diatesi aspettuale: mentre la seconda voce è nella fase ascendente del suo soggetto (accomplishment in corso), la prima voce è nella nota tenuta della sua risposta (stato), creando la prima asimmetria temporale del brano — una voce "in moto" e una voce "ferma" nello stesso istante.

Questa asimmetria è gestita da Bach con una precisione che sembra pianificata per massimizzare la tensione aspettuale. Le misure cinque-otto mostrano la terza entrata del soggetto (voce di alto) mentre le prime due voci sono già nell'area degli episodi liberi: la voce di basso è in un'attività atelica (una figurazione di semicrome discendenti che non tende verso nessun terminus specifico), la voce di tenore è in uno stato (un'armonia tenuta sull'accordo di dominante). La terza entrata del soggetto introduce un accomplishment in questa texture di stato più attività, creando una dissincronia aspettuale tripla: stato (tenore) più attività (basso) più accomplishment in corso (alto). Il tempo musicale di queste misure è qualitativamente plurale: si è simultaneamente "fermi", "in moto atelico" e "in cammino verso un terminus", e questa simultaneità è la fonte di una ricchezza temporale che nessuna delle tre voci produrrebbe da sola.

Le misure dodici-sedici mostrano la configurazione più interessante dell'intera fuga: il momento in cui tutte e tre le voci si trovano simultaneamente in diversi stati di un ciclo cadenzale. La voce superiore ha appena completato un breve accomplishment (una frase melodica con cadenza locale — achievement della risoluzione cadenzale), la voce mediana è nel mezzo di un'attività (una sequenza cromatica discendente), la voce inferiore sta iniziando un nuovo accomplishment (l'inizio di una nuova entrata del soggetto nella tonalità relativa). Il "post-achievement" del soprano — lo stato conseguente all'achievement della cadenza, la nota tenuta che segue la risoluzione — è

una delle categorie aspettuali più sottili e più musicalmente rilevanti: è lo stato che risulta dal completamento di un evento telico, analogo al *present perfect* risultativo inglese. La sua coesistenza con un'attività in corso e un nuovo accomplishment che si sta aprendo è precisamente il tipo di tensione aspettuale che Bach costruisce con notevole consapevolezza strutturale.

Le misure trenta-trentasei mostrano il climax aspettuale della fuga: la stretto, in cui le entrate del soggetto si sovrappongono a distanza di una misura. In questa sezione, tutte e tre le voci sono simultaneamente in accomplishment, ma con sfasamenti temporali tali che i rispettivi terminus non coincidono mai. La voce superiore è all'achievement del suo terminus quando la voce mediana è nel mezzo del suo accomplishment quando la voce inferiore è appena all'inizio del suo accomplishment. Questa configurazione — tre accomplishments sfasati che raggiungono i loro terminus in momenti diversi — produce la massima tensione aspettuale possibile: non c'è mai un momento in cui tutte le voci sono simultaneamente in stato di post-achievement (il momento di "riposo" post-terminale), e quindi l'ascoltatore è tenuto costantemente in stato di attesa di completamenti che arrivano uno alla volta, sempre prima che il precedente sia pienamente "assorbito". Questo è il meccanismo aspettuale della stretto bachiana: non solo la compressione temporale delle entrate del soggetto (che è la descrizione formale standard) ma la produzione di una tensione aspettuale di massima intensità attraverso la sovrapposizione di accomplishments sfasati che non si lasciano mai sincronizzare.

La risoluzione della stretto — le misure conclusive in cui la fuga converge sulla cadenza finale — è la sincronizzazione aspettuale: il momento in cui tutte e tre le voci convergono sullo stesso regime temporale (l'accomplishment della cadenza autentica finale) e raggiungono simultaneamente lo stesso terminus (la tonica di do diesis minore). Questa sincronizzazione finale è l'evento di massima soddisfazione aspettuale del brano: non solo la risoluzione armonica sulla tonica ma la cessazione della tensione aspettuale plurale, il collasso del tempo multiplo in un unico punto di arrivo condiviso. La corrispondenza di tutti e tre i terminus nello stesso istante è la struttura cognitiva della "chiusura" che ogni ascoltatore percepisce alla fine di una fuga ben costruita.

4.4 La coercion aspettuale in musica: quando il contesto forza il cambiamento di tipo

Il fenomeno della coercion aspettuale — introdotto brevemente nel capitolo precedente a proposito di Coltrane — merita uno sviluppo sistematico perché è uno dei meccanismi generativi più importanti del contrappunto avanzato e della composizione musicale sofisticata. In linguistica, la coercion si verifica quando un verbo con una certa Aktionsart lessicale è usato in un contesto grammaticale che richiede una Aktionsart diversa. Il caso classico è quello dei verbi di stato usati nel progressivo inglese: "she is being difficult" forza il verbo stativo "to be" in un'interpretazione accomplishment-like, suggerendo che l'essere-difficile è un comportamento intenzionale e temporaneo. Il verbo non cambia morfologia, ma il

contesto grammaticale lo forza a sviluppare un'interpretazione temporale diversa da quella lessicale.

In musica, la coercion aspettuale si verifica quando una struttura musicale con una certa Aktionsart intrinseca è collocata in un contesto che forza l'ascoltatore a reinterpretarla secondo una Aktionsart diversa. Il caso più comune è la coercion telicizzante: una struttura atelica che, collocata in un contesto di forte aspettativa cadenzale, viene percepita come un accomplishment in corso. Un ostinato di basso che si ripete su una dominante prolungata è intrinsecamente atelico (potrebbe continuare indefinitamente), ma quando è collocato alla fine di uno sviluppo sonatistico — dove l'aspettativa della ripresa è massimalmente intensa — viene percepito come un accomplishment che tende verso il terminus della ripresa stessa. Il contesto formale coerisce l'atelicità dell'ostinato in una telicità contestuale: la struttura sonora non cambia, ma la sua Aktionsart cambia perché il contesto impone una diversa struttura temporale.

Il caso inverso — la coercion atelicizzante, in cui un evento telico viene forzato in un'interpretazione atelica dal contesto — è meno comune ma altrettanto reale. La ripetizione insistita di una cadenza autentica — come nelle misure conclusive di molte sinfonie classiche, dove la cadenza finale viene ripetuta quattro, sei, otto volte — produce una coercion atelicizzante: la prima cadenza autentica è un accomplishment che raggiunge il suo terminus; la seconda ripetizione trasforma retroattivamente la prima in parte di un pattern iterativo atelico, e da quel momento la cadenza non è più percepita come terminus ma come unità di un'attività abituale. Questo meccanismo spiega perché le conclusioni delle sinfonie classiche producono effettivamente un senso di meta-commento sulla telicità stessa: le cadenze finali si auto-commentano come cadenze, trasformando la telicità in atelicità attraverso l'iterazione.

Beethoven è il compositore che ha esplorato più sistematicamente questo meccanismo nelle conclusioni. Le misure conclusive della Quinta Sinfonia (1808) ripetono la cadenza finale in do maggiore con una frequenza che inizialmente sorprende — si contano, nelle ultime cinquanta misure, almeno diciotto cadenze autentiche o mezze autentiche sulla tonica di do maggiore — ma che ha una funzione aspettuale precisa: trasformare progressivamente la telicità della cadenza in un gesto affermativo e perentorio che ha perduto la funzione di terminus e ha assunto la funzione di dichiarazione atelica. Le ultime misure della Quinta non "arrivano alla tonica": sono già "nella tonica", la abitano come stato, la dichiarano come qualità stabile del mondo sonoro. La coercion atelicizzante ha trasformato l'accomplishment in attività-dichiarativa, e questa trasformazione — che si svolge nell'arco delle ultime cinquanta misure — è l'evento aspettuale più grande dell'intera sinfonia, più grande di qualsiasi singolo evento armonico o ritmico.

4.5 La teoria di Hasty riletta attraverso la diatesi aspettuale

Christopher Hasty, nell'analisi del tempo musicale come processo di proiezione descritta nel Capitolo I, ha sviluppato senza nominarla una teoria implicita della diatesi aspettUALE. La nozione hastyiana di "anacrusis" è precisamente la struttura dell'accomplishment in corso nel senso di Vendler: un evento che ha già aperto il suo processo verso un terminus e che "porta avanti" la sua direzionalità verso la downbeat. La downbeat è il terminus dell'accomplishment anakrusitico, il suo achievement terminale. La "continuation" che segue la downbeat è lo stato post-achievement: il tempo si è assestato dopo la culminazione, e la voce è in un "presente esteso" che non tende ancora verso nessun nuovo terminus.

La teoria di Hasty acquista una nuova dimensione quando la si legge attraverso la lente della diatesi aspettUALE polifonica. Hasty analizza principalmente il metro come fenomeno che si svolge su una sola linea temporale — come se il tempo musicale fosse sempre monodico. Ma nella polifonia, la proiezione metrica si svolge simultaneamente su più linee, e le diverse voci possono trovarsi in fasi diverse del ciclo anacrusis-downbeat-continuation. La tensione che Hasty chiama "incompleteness" — la sensazione che un evento metrico sia "aperto", in attesa del suo completamento — è, nella prospettiva di questa monografia, la tensione della diatesi aspettUALE: la voce che è in anacrusis è in un accomplishment in corso, la voce che è in downbeat è all'achievement del suo terminus, la voce che è in continuation è nel post-achievement stativo.

Questa rilettura di Hasty attraverso l'aspetto ha una conseguenza teorica importante: mostra che la teoria del metro e la teoria dell'aspetto non sono teorie parallele che descrivono dimensioni diverse della musica, ma teorie che descrivono la stessa struttura cognitiva da prospettive diverse. Il metro è l'aspetto applicato al livello della pulsazione ritmica; l'aspetto è il metro applicato al livello della struttura armonica. Questa convergenza suggerisce l'esistenza di una teoria più generale del "tempo interno degli eventi" che unifica la teoria aspettUALE e la teoria metrica in un unico framework — una possibilità che si indica qui come programma di ricerca futura.

4.6 Diatesi aspettUALE e tensione narrativa: la forma come sequenza di configurazioni

L'ultima tesi di questo capitolo — e la più ambiziosa nelle implicazioni — è che la forma musicale può essere descritta come una sequenza di configurazioni di diatesi aspettUALE, e che la "narrativa" di un brano — il suo arco drammatico, la sua struttura di tensione e risoluzione — è almeno parzialmente costituita dalla sequenza di queste configurazioni. Questa tesi integra e supera la descrizione puramente formale della forma musicale e la descrizione puramente armonica: propone che la forma abbia anche una dimensione aspettUALE, che i cambiamenti nella configurazione di diatesi siano eventi narrativi che producono variazioni nella qualità del tempo musicale che l'ascoltatore percepisce come "accadimenti" nell'arco del brano.

La forma-sonata, riletta attraverso la diatesi aspettUALE, mostra una struttura narrativa che la descrizione formale tradizionale non cattura interamente. L'esposizione è caratterizzata da alta sincronizzazione aspettUALE: le voci si muovono in configurazioni di diatesi relativamente uniformi, con pochi momenti di forte dissincronia, producendo un tempo musicale che tende alla chiarezza e alla prevedibilità. Lo sviluppo è caratterizzato da massima dissincronia aspettUALE: le voci si trovano frequentemente in stati aspettuali diversi, le sequenze ateliche si mescolano con accomplishments interrotti, le zone statiche si alternano con zone di massima direzionalità telica, producendo un tempo musicale plurale e instabile. La ripresa è caratterizzata dalla progressiva re-sincronizzazione aspettUALE: le voci convergono verso configurazioni sempre più uniformi, fino alla sincronizzazione finale della cadenza conclusiva.

Questa narrativa aspettUALE — sincronizzazione, poi dissincronia, poi re-sincronizzazione — è isomorfa alla narrativa armonica della forma-sonata (stabilità, poi modulazione, poi ritorno) ma non è la stessa cosa: si può avere la narrativa armonica senza la narrativa aspettUALE (un brano che modula senza produrre dissincronia aspettUALE) e viceversa (un brano che non modula ma produce forte dissincronia aspettUALE attraverso la sovrapposizione di voci in regimi temporali diversi). L'indipendenza parziale delle due narrative è la prova che sono dimensioni distinte della struttura musicale, e che una descrizione completa della forma richiede entrambe.

5. Le Tradizioni Ritmiche Non-Occidentali

Terreno di verifica e fonte di tipologie aspettuali nuove

5.1 Perché le tradizioni non-occidentali

L'analisi delle tradizioni musicali sviluppata nel Capitolo III ha il limite implicito di muoversi all'interno di un orizzonte — quello della musica colta occidentale e delle sue derivazioni novecentesche, compreso il jazz — che condivide presupposti cognitivi e strutturali sufficientemente omogenei da rendere la tipologia aspettUALE relativamente facile da costruire. I sistemi analizzati sono tutti, in modi diversi, eredi o reazioni a una tradizione che ha sviluppato la grammatica della telicità armonica come suo asse principale: anche il minimalismo, nella sua ribellione alla telicità classica, si definisce in relazione a essa. Per verificare che la teoria aspettUALE non sia semplicemente una proiezione di questa tradizione su tutte le tradizioni musicali — per testare se le

categorie vendleriane siano genuinamente universali o siano invece specifiche di un certo tipo di organizzazione del tempo musicale — è necessario uscire da questo orizzonte e confrontarsi con sistemi musicali che hanno sviluppato la loro organizzazione del tempo in modo indipendente e strutturalmente diverso.

Tre tradizioni sono scelte per questa analisi comparativa: il sistema del *Tala* nella musica classica dell'India del nord (musica indostanica), il gamelan giavanese e balinese come sistema ritmico-timbrico senza la centralità dell'armonia funzionale, e le polifonie vocali dell'Africa subsahariana come terreno in cui la struttura aspettatale è realizzata attraverso l'intreccio di pattern ritmici indipendenti piuttosto che attraverso la progressione armonica. Ciascuna di queste tradizioni è presentata con la consapevolezza dei suoi limiti: l'analisi aspettatale condotta qui è necessariamente parziale e preliminare, costruita su fonti secondarie e su ascolti analitici che non possono sostituire la conoscenza profonda che solo l'immersione prolungata nelle tradizioni stesse produce. L'obiettivo non è l'eshaustività etnomusicologica ma la messa alla prova della teoria.

5.2 Il *Tala* indostanico: accomplishments ciclici e la telicità della somma

Il *Tala* — il sistema di organizzazione temporale della musica classica dell'India del nord, condiviso nelle sue linee fondamentali anche dalla tradizione del sud — è uno dei sistemi di organizzazione ritmica più elaborati e più formalmente descritti che l'etnomusicologia conosca. Un *tala* è un ciclo ritmico di durata fissa, articolato internamente in sezioni di diversa lunghezza (*vibhag*), ciascuna delle quali è caratterizzata da un gesto di battuta specifico (*theke*) che il suonatore di *tablā* esegue come marcatore strutturale. Il ciclo si ripete indefinitamente nel corso di un'esecuzione, e la sua durata — da sei battiti del *Dadra tala* ai sedici dell'*Adital* o ai trentadue misure dell'*Adha Chowtala* nei cicli più estesi — determina il "respiro" ritmico fondamentale dell'esecuzione.

La questione aspettatale fondamentale posta dal sistema del *tala* è la seguente: il ciclo del *tala* è telico o atelico? In apparenza, la risposta sembra semplice: il ciclo si ripete, ogni ripetizione è strutturalmente equivalente alle altre, e il "fine" dell'esecuzione è sempre una decisione pratica esterna alla struttura del *tala* — caratteristiche tutte di un'attività atelica. Ma questa risposta, pur corretta a livello macroformale, ignora una struttura aspettatale interna al ciclo che non ha equivalenti nel sistema tonale occidentale.

All'interno di ogni singolo ciclo del *tala*, la struttura aspettatale è quella di un accomplishment: le sezioni del ciclo si succedono in un ordine fisso, ogni sezione prepara la successiva, e il *sam* — il primo battito del ciclo, il momento di massimo peso strutturale e accentuale — è il terminus dell'accomplishment intracyclico. Il *sam* non è un termine qualsiasi: è il punto di risoluzione verso cui l'intera architettura del ciclo è orientata, il momento in cui l'improvvisazione melodica (eseguita dal solista) e

la base ritmica (eseguita dal *tablā*) si incontrano e si "resincronizzano" dopo il libero dispiegarsi dell'improvvisazione nello spazio del ciclo. Questo incontro sul *sam* è il *terminus* intrinseco del ciclo: non è arbitrario né intercambiabile con qualsiasi altro momento del ciclo (l'irriducibilità dell'*endpoint*), produce una qualità di "arrivo" percettivamente marcata e cognitivamente saliente (la direzionalità interna), e la sua mancanza — quando il solista "non arriva" al *sam* nel momento previsto, o lo supera per errore — produce un'incoerenza strutturale percepibile e formalmente stigmatizzata nella pratica performativa (la sospensibilità senza incoerenza).

La struttura aspettuale del *tala* è dunque quella di un *accomplishment* iterato: ogni ciclo è un *accomplishment* con *terminus* intrinseco sul *sam*, e la successione dei cicli è un'attività abituale che replica indefinitamente la struttura dell'*accomplishment* singolo senza che nessuna iterazione sia intrinsecamente più "completa" di un'altra. Questa struttura — *accomplishment* iterato in un'attività abituale — non ha equivalenti precisi nella musica tonale occidentale, dove l'iterazione delle strutture teliche (come la ripetizione della cadenza finale nelle sinfonie beethoveniane) ha l'effetto di atelicizzare la struttura, di toglierle la qualità di *terminus* e conferirle la qualità di dichiarazione stativa. Nel sistema del *tala*, l'iterazione non atelicizza: ogni arrivo al *sam* mantiene pienamente la sua qualità di *terminus*, pienamente la sua qualità di *achievement*, indipendentemente da quante volte il ciclo si è già ripetuto. Questo produce un regime aspettuale che non ha nome nella tipologia classica: né la telicità-una-volta del sistema classico occidentale, né l'atelicità dell'ostinato minimalista, ma una telicità-ciclica che è al tempo stesso sempre la stessa e sempre nuova.

La pratica dell'improvvisazione nel sistema indostanico illumina ulteriormente questa struttura aspettuale peculiare. L'improvvisatore non "crea" il *terminus* — il *sam* esiste indipendentemente da lui, è dato dal *tala* — ma "gioca" con la distanza e l'avvicinamento al *terminus*: può allontanarsene nelle variazioni di lunga durata (*tihais*), può fintare di arrivarci prima del tempo con cadenze che simulano il *sam* senza coincidere con esso, può caricare il percorso verso il *sam* di densità ritmica crescente che aumenta la direzionalità aspettuale dell'*accomplishment* intracyclico. In questo senso, il musicista indostanico pratica una forma sofisticata di manipolazione della telicità: sa con esattezza quando sarà il *terminus*, sa esattamente quale sarà il *terminus*, e questa certezza — radicalmente diversa dall'incertezza che caratterizza l'esperienza della cadenza nella musica occidentale, dove il *terminus* è atteso ma non è mai del tutto prevedibile nei dettagli della sua realizzazione — è il presupposto cognitivo che rende possibile l'elaborazione ritmica della distanza dal *terminus* stesso.

5.3 Il Gamelan giavanese: stati prolungati e l'aspetto stratificato delle texture

Il gamelan giavanese — l'orchestra di percussioni metalliche (*gong*, *saron*, *bonang*, *kenong*, *kempul*) che costituisce il nucleo della vita musicale cortese e cerimoniale di Giava — presenta un sistema di organizzazione temporale che sfida in modo particolarmente produttivo le categorie aspettuative sviluppate per la musica tonale

occidentale. Il gamelan non ha armonia funzionale nel senso europeo: non ci sono accordi che "puntano" verso una tonica, non ci sono progressioni cadenzali che creino direzionalità nel senso che questa monografia ha analizzato nel sistema classico. La struttura temporale del gamelan è organizzata intorno a principi diversi: la stratificazione di pattern ritmici di diversa scala temporale, la periodicità gerarchica dei diversi strumenti, e la funzione dei grandi gong come marcatori strutturali di livelli diversi della forma.

Nella struttura del gamelan, i diversi strumenti suonano pattern ritmici di diversa lunghezza che si sovrappongono: il *saron* esegue la melodia principale in valori regolari e relativamente veloci; i *bonang* ornamentano e anticipano la melodia con figurazioni più dense; il *kempul* e il *kenong* suonano a intervalli più radi, marcando le sezioni intermedie della forma; il grande gong (*gong ageng*) suona solo all'inizio e alla fine di ogni ciclo formale completo, che può durare decine di minuti. Questa stratificazione — dal livello più denso e rapido del *saron* al livello più rarefatto e lento del gong — è una stratificazione aspettuale nel senso che i diversi strumenti operano a livelli temporali diversi e producono strutture aspettuali diverse che coesistono simultaneamente.

Il *saron*, con la sua melodia omogenea e continua, produce un'attività aspettuale: il processo è dinamico, si svolge nel tempo, ma non tende verso nessun terminus specifico nel breve periodo — ogni nota è del medesimo tipo delle altre, ogni segmento della melodia potrebbe continuare indefinitamente senza che la struttura cambi tipo. Il *kempul*, con i suoi colpi periodici che marcano le sezioni, produce accomplishments locali di piccola scala: ogni colpo è il terminus di un breve processo di attesa, un achievement di scala locale. Il *gong ageng*, con la sua rarità e il suo peso strutturale enorme, produce l'achievement di massima scala: il momento in cui arriva il gong è il terminus dell'intero ciclo formale, l'arrivo al *sam* del gamelan, per così dire.

La diatesi aspettuale nel gamelan è dunque una diatesi stratificata non per voci contrappuntistiche — come in Bach, dove ogni voce è una linea melodica indipendente — ma per livelli temporali dello stesso tessuto sonoro. Ogni strumento del gamelan abita un livello temporale diverso, e la loro coesistenza produce una pluralità di regimi temporali simultanei che non è riducibile a nessuno di essi preso singolarmente. Un ascoltatore che ascolti solo il *saron* sentirebbe un'attività omogenea e potenzialmente infinita; un ascoltatore che ascolti solo il gong sentirebbe la cadenza di un lunghissimo accomplishment che si completa nell'achievement del colpo finale; un ascoltatore che ascolti il gamelan nella sua totalità sperimenta la coesistenza di questi regimi temporali in una diatesi aspettuale di tipo inedito: non la dissincronia tra voci contrappuntistiche dello stesso livello temporale, come in Bach, ma la simultaneità di livelli temporali di scale radicalmente diverse, da cui emerge una qualità temporale che non è né puramente atelica né puramente telica ma qualcosa di terzo.

Questa struttura aspettuale stratificata per livelli temporali è stata descritta dalla musicologia del gamelan — in particolare da Mantle Hood (1954) e dalla tradizione di

analisi sviluppata da Marc Perlman — in termini di "gongan" (la sezione delimitata dai grandi gong), di "colotomic structure" (la struttura di marcatura temporale gerarchica) e di "balungan" (la melodia ossatura su cui si costruisce l'ornamentazione degli altri strumenti). Ma la descrizione aspettuale permette di vedere in questi concetti qualcosa che la terminologia musicologica tradizionale non rende esplicito: il fatto che la coesistenza di livelli temporali diversi non è semplicemente una questione di metro o di densità ritmica, ma una questione di qualità aspettuale — di diversi regimi del tempo interno che abitano lo stesso spazio sonoro senza annullarsi a vicenda.

5.4 Le polifonie vocali dell'Africa subsahariana: aspetto e pattern interlocking

Le polifonie vocali e strumentali dell'Africa subsahariana — dalle polifonie dei Ba-Benzélé della Repubblica Centrafricana alle polifonie vocali dei Venda del Sudafrica, dalle musiche per xilofono dei Chopi del Mozambico ai pattern interlocking degli ensemble di flauti e corni dei pastori dell'Africa orientale — presentano un sistema di organizzazione del tempo che la teoria aspettuale può descrivere in modo nuovo rispetto alla musicologia tradizionale.

Il meccanismo strutturale fondamentale di molte di queste tradizioni è quello dell'*interlocking* — l'intreccio di parti musicali brevi, ognuna delle quali è incompleta da sola, che si combinano per produrre un pattern ritmico-melodico complesso che nessuna singola voce contiene interamente. Questo meccanismo — descritto da Simha Arom (1985) con il termine di "hocket" poliritmico — ha una struttura aspettuale peculiare che non ha equivalenti precisi nelle tradizioni analizzate nei capitoli precedenti.

Ogni singola parte di un ensemble di interlocking è, presa isolatamente, strutturalmente incompleta nel senso aspettuale più preciso: non è semplicemente "breve" ma è strutturata in modo tale da non avere senso compiuto senza le altre parti. Un flautista che suona la sua parte in uno schema di interlocking a tre non sta suonando un'attività autonoma (che potrebbe continuare indefinitamente e avere senso da sola) né un accomplishment autonomo (che tenderebbe verso un terminus proprio): sta suonando un frammento che "aspetta" le altre parti per completarsi, che ha la sua struttura temporale interna ma la realizza solo nell'intreccio con le parti complementari. Questo "aspettare le altre parti" non è una metafora intenzionalista: è una struttura cognitiva verificabile, nel senso che l'interruzione di una parte sola — senza le altre — produce esattamente la stessa qualità di incoerenza che si produce nell'interruzione di un accomplishment prima del suo terminus.

La struttura aspettuale dell'interlocking africano è dunque quella di un accomplishment distribuito: un processo che ha un terminus intrinseco, che è direzionato verso di esso, e la cui struttura interna è distribuita tra voci diverse che contribuiscono ciascuna una porzione del processo totale. Il terminus di questo accomplishment distribuito non è raggiunto da nessuna singola voce ma dall'insieme

delle voci nel momento in cui i loro pattern si combinano per completare il ciclo. Questa struttura — in cui il terminus è una proprietà emergente dell'insieme piuttosto che una proprietà di qualsiasi singola voce — non ha equivalenti nel contrappunto occidentale, dove ogni voce ha la propria struttura aspettUALE (il suo accomplishment, la sua attività, il suo stato) e la diatesi aspettUALE è la coesistenza di strutture autonome. Nell'interlocking africano, la struttura aspettUALE è costitutivamente distribuita: nessuna voce "ha" da sola un accomplishment completo, ma l'insieme "ha" l'accomplishment che nessuna voce singola possiede.

5.5 Implicazioni teoriche comparative: tre tipologie aspettuali nuove

Il confronto tra le tre tradizioni analizzate in questo capitolo e le tradizioni occidentali analizzate nel Capitolo III rivela tre tipologie aspettuali che non erano prevedibili dall'interno della sola prospettiva occidentale e che arricchiscono significativamente la teoria:

La prima è la *telicità ciclica* del tala indostanico: un sistema in cui il terminus è intrinseco al ciclo (e quindi l'aspetto è telico) ma l'iterazione del ciclo non atelicizza la telicità (come invece succede nella coercion atelicizzante del classicismo viennese quando le cadenze si ripetono). La telicità ciclica produce un regime temporale in cui ogni arrivo è genuinamente un terminus e ogni terminus è genuinamente un achievement, indipendentemente dal numero di iterazioni precedenti. Questo tipo di rapporto con il terminus — che si potrebbe chiamare "telicità dell'eterna novità" — è cognitivamente diverso dalla telicità-una-volta del sistema classico e dall'atelicità iterativa del minimalismo.

La seconda è la *stratificazione aspettUALE per scala temporale* del gamelan giavanese: un sistema in cui la diatesi aspettUALE non è tra voci dello stesso livello temporale (come nella polifonia bachiana) ma tra livelli temporali radicalmente diversi che coesistono nella stessa texture sonora. In questo sistema, non si percepisce una "tensione" tra diatesi diverse nel senso della dissincronia bachiana — la differenza di scala temporale è troppo grande per produrre tensione in senso stretto — ma una qualità di pluralità temporale stratificata, in cui diversi "orizzonti di completamento" coesistono e si sovrappongono senza interferirsi.

La terza è la *telicità distribuita* delle polifonie di interlocking africano: un sistema in cui il terminus non appartiene a nessuna voce individuale ma è una proprietà emergente dell'ensemble. In questo sistema, la diatesi aspettUALE non è tra voci con Aktionsart diverse ma tra frammenti aspettualmente incompleti che si completano reciprocamente nell'intreccio.

Queste tre tipologie non contraddicono il framework aspettUALE sviluppato nei capitoli precedenti: mostrano che esso è capace di accogliere strutture aspettuali che vanno oltre il repertorio della musica tonale occidentale e che le categorie fondamentali —

stato, attività, accomplishment, achievement, telicità, atelicità — hanno risorse teoriche ancora inesplorate che la sola prospettiva occidentale non poteva rivelare. Il framework attuale è genuinamente crossculturale non perché imponga le sue categorie a tutte le tradizioni in modo uguale, ma perché quelle categorie sono sufficientemente flessibili e sufficientemente generali da accogliere variazioni che modificano le relazioni tra le categorie stesse — la relazione tra iterazione e telicità, la relazione tra scala temporale e diatesi, la relazione tra voce individuale e terminus collettivo — senza dissolvere la distinzione fondamentale tra eventi con terminus intrinseco ed eventi senza terminus intrinseco.

5.6 Limiti e cautele metodologiche

Prima di procedere alle conclusioni, è necessario nominare esplicitamente i limiti dell'analisi comparativa sviluppata in questo capitolo. Il più importante è quello della mediazione: l'analisi del tala, del gamelan e delle polifonie africane è condotta su fonti secondarie e su ascolti analitici di registrazioni, non su conoscenza diretta delle tradizioni sviluppata attraverso l'immersione e la pratica. Questo limite non invalida l'analisi ma ne circoscrive il raggio: le strutture aspettuali identificate sono plausibili e teoricamente coerenti, ma richiederebbero la verifica di etnomusicologi e di musicisti appartenenti alle tradizioni in questione per essere confermate o corrette nei dettagli.

Il secondo limite è quello della selettività: le tre tradizioni scelte — tala indostanico, gamelan giavanese, polifonie africane — sono rappresentative di aree geografiche e culturali vastissime all'interno delle quali esistono differenze enormi che questa analisi non cattura. "Musica classica nordindiana" è un'etichetta che copre tradizioni regionali, scuole (*gharana*), strumentazioni e pratiche performative profondamente diverse; "gamelan" è un termine che abbraccia almeno due tradizioni principali (giavanese e balinese) che differiscono significativamente nella struttura; "polifonie africane" è una categoria così ampia da essere quasi priva di senso se non specificata in modo molto più granulare. La scelta di lavorare a questo livello di astrazione è una necessità metodologica di questo saggio, non una scelta intellettuale ideale, e il limite che ne risulta deve essere tenuto presente nella valutazione delle conclusioni.

C. Il Tempo Interno della Musica

Bilancio di una proposta

C.1 Che cosa è stato dimostrato

Questo saggio è partito da una domanda che la teoria musicale non aveva ancora posto con questa precisione: la musica ha aspetto? Non nel senso vago di avere

"dimensioni temporali", che nessuno avrebbe negato, ma nel senso preciso e operativo in cui la linguistica usa questo termine — nel senso di una struttura cognitiva che organizza la qualità temporale interna degli eventi come telici o atelici, stativi o dinamici, puntuali o duraturi. La risposta proposta in questi cinque capitoli è affermativa, ed è stata argomentata a tre livelli distinti che vale la pena ricapitolare nella loro interrelazione.

Il primo livello è quello della tassonomia degli eventi musicali. Il Capitolo II ha mostrato che le quattro categorie di Aktionsart di Vendler — stati, attività, accomplishments, achievement — si applicano alla musica con criteri operazionali precisi e verificabili: la sospendibilità senza incoerenza, la cumulatività, la direzionalità interna, l'irriducibilità dell'endpoint. Questi criteri producono distinzioni che non sono arbitrarie ma cognitivamente fondate: la differenza tra un pedale di tonica (stato) e una progressione cadenzale (accomplishment) non è solo una differenza di struttura armonica ma una differenza nella qualità del tempo che producono, nel tipo di attesa che generano, nella fenomenologia dell'ascolto. La tassonomia aspettuale degli eventi musicali non è un sistema di etichette importate dalla linguistica: è la descrizione di strutture cognitive che la musica ha sempre prodotto e che la teoria non aveva ancora nominato. L'analisi della musica vocale come dominio in cui due strati aspettuativi — quello del testo e quello della composizione — si sovrappongono, si sincronizzano e talvolta confliggono ha mostrato che il controaspetto compositivo (la scelta deliberata di contrastare l'aspetto del testo con l'aspetto della musica) è uno degli strumenti espressivi più sottili a disposizione dei compositori, usato con consapevolezza magistrale da Mozart, Schubert e Wolf.

Il secondo livello è quello della tipologia delle tradizioni. Il Capitolo III ha mostrato che le diverse tradizioni musicali grammaticalizzano l'aspetto in modi sistematicamente diversi: il classicismo viennese grammaticalizza la telicità attraverso la gerarchia cadenzale obbligatoria; il minimalismo americano grammaticalizza l'atelicità attraverso la rimozione strutturale del terminus come categoria organizzativa; il jazz bebop grammaticalizza la densità di accomplishments rapidi attraverso la progressione accordale obbligatoria; il contrappunto rinascimentale lascia l'aspetto relativamente non grammaticalizzato nel tessuto interstiziale, marcandolo solo nei momenti cadenzali; la musica spettrale grammaticalizza l'aspetto attraverso la dimensione timbrica piuttosto che quella armonica. Questa variazione tipologica non è accidentale né puramente stilistica: è la manifestazione di diverse soluzioni culturalmente sviluppate a un problema cognitivo condiviso.

Il terzo livello — e il più originale di questo saggio — è quello della diatesi aspettuale nella polifonia. Il Capitolo IV ha mostrato che la simultaneità delle voci contrappuntistiche produce configurazioni di diatesi aspettuale — coesistenze di stati aspettuativi diversi nello stesso spazio temporale — che sono fonti di una tensione cognitiva specifica, irriducibile alla tensione armonica e alla tensione ritmica. L'analisi sistematica della fuga in do diesis minore di Bach ha mostrato come queste configurazioni si succedano nel corso del brano in una sequenza che costituisce una

narrativa aspettuale autonoma, isomorfa ma non identica alla narrativa armonica. La coercion aspettuale — il meccanismo attraverso cui il contesto forza eventi in Aktionsart diverse dalle loro proprietà intrinseche — è uno degli strumenti compositivi più potenti della tradizione contrappuntistica e romantica, e la sua descrizione attraverso la teoria aspettuale ne rivela la struttura cognitiva con una precisione che la sola teoria armonica non poteva raggiungere.

Il quarto livello, introdotto dal Capitolo V, è quello della verifica crossculturale. L'analisi del tala indostanico, del gamelan giavanese e delle polifonie africane di interlocking ha mostrato che il framework aspettuale è capace di accogliere strutture che non erano prevedibili dall'interno della sola prospettiva occidentale — la telicità ciclica, la stratificazione aspettuale per scala temporale, la telicità distribuita — e che queste strutture non contraddicono il framework ma lo arricchiscono, mostrando che le relazioni tra le categorie fondamentali sono più flessibili di quanto l'analisi della musica occidentale avrebbe suggerito.

C.2 Il rapporto con la semiotica musicale e la teoria narrativa

Un punto di collegamento che questo saggio ha deliberatamente lasciato in ombra — per non sovraccaricare un'argomentazione già densa — merita di essere tematizzato nelle conclusioni: il rapporto tra la teoria aspettuale degli eventi musicali e le tradizioni della semiotica musicale e della teoria narrativa applicate alla musica, rappresentate soprattutto da Eero Tarasti (1994), da Jean-Jacques Nattiez (1987) e, in una prospettiva diversa, da Lawrence Zbikowski (2002).

Tarasti, nella sua semiotica musicale ispirata a Greimas, ha proposto che la musica possa essere analizzata come un sistema di eventi dotati di "isotopie" — qualità semantiche che si ripetono e si trasformano nel corso del brano — e ha introdotto il concetto di "modalità" musicale (sapere, potere, volere, dover fare) come dimensioni del significato musicale. Questa proposta converge con la teoria aspettuale nel riconoscere che gli eventi musicali hanno qualità cognitive che vanno al di là della loro descrizione acustica, ma diverge nel suo presupposto che queste qualità siano di natura semiologica — che la musica "significhi" nel senso greimasiano — piuttosto che cognitiva-strutturale nel senso proposto qui. La differenza non è sottile: per Tarasti, le qualità temporali degli eventi musicali sono funzione del loro significato semantico (un evento "vuole" arrivare da qualche parte perché è semanticamente investito di desiderio); per questa monografia, le qualità temporali sono strutture cognitive che esistono indipendentemente da qualsiasi significato semantico proiettato sull'evento.

Nattiez, con la sua distinzione tra livello poietico (il processo compositivo), *esthésique* (il processo percettivo) e neutro (la struttura fisica del testo musicale), ha fornito il framework metodologico più influente per l'analisi semiotica della musica. La teoria aspettuale si colloca prevalentemente a livello *esthésique* nel senso di Nattiez: riguarda la qualità cognitiva dell'esperienza percettiva degli eventi musicali, non la loro struttura fisica né le intenzioni del compositore. Ma questa collocazione è parziale: la

grammaticalizzazione dell'aspetto nelle tradizioni musicali — la gerarchia cadenzale del classicismo viennese, le strutture iterative del minimalismo — è anche un fatto del livello neutro, una proprietà strutturale del testo che produce effetti cognitivi prevedibili nell'ascolto.

Zbikowski, nella sua proposta di una "musica come struttura concettuale" ispirata a Lakoff e Johnson, ha sostenuto che la comprensione musicale è fondamentalmente metaforica: l'ascoltatore proietta strutture concettuali del dominio fisico (spazio, moto, forza) sul dominio musicale (altezza, progressione, tensione). Questa proposta è la più vicina alla teoria aspettuale tra quelle della tradizione semiotica, perché condivide il presupposto che la comprensione musicale operi attraverso strutture cognitive generali che si applicano a domini diversi. Ma la teoria aspettuale va oltre la proposta di Zbikowski in un senso preciso: non si limita a sostenere che la musica sia compresa "metaforicamente" attraverso categorie fisico-spaziali, ma sostiene che la struttura aspettuale — la distinzione telico/atelico, la qualità della direzionalità verso un terminus, la struttura del cambiamento puntuale — è una struttura cognitiva propria che non è derivata per metafora dal dominio fisico ma condivisa tra musica e linguaggio come categoria cognitiva genuinamente comune.

C.3 Ciò che resta aperto

Come in ogni ricerca seria, le questioni aperte sono più numerose e più interessanti di quelle chiuse. Tre meritano di essere nominate esplicitamente come priorità per la ricerca futura.

La prima è la verifica sperimentale della diatesi aspettuale. La tesi che la dissincronia aspettuale produce una tensione cognitiva distinta dalla tensione armonica e dalla tensione ritmica è una tesi empiricamente verificabile: si possono costruire stimoli musicali con la stessa struttura armonica e ritmica ma diverse configurazioni di diatesi, e misurare le risposte degli ascoltatori — tempi di reazione, potenziali evocati, misure di tensione soggettiva — per verificare se le diverse configurazioni producono effetti distinguibili. Questa verifica richiede una collaborazione tra teorici musicali, psicologi cognitivi e neuroscienziati che non è ancora stata condotta, e la sua conduzione richiederebbe la costruzione di stimoli controllati sufficientemente complessi da realizzare configurazioni di diatesi aspettuale genuine.

La seconda è la relazione tra aspetto musicale e aspetto linguistico nello sviluppo cognitivo. Se la struttura aspettuale è una categoria cognitiva condivisa tra musica e linguaggio, si dovrebbe osservare una correlazione tra la precocità dell'acquisizione dell'aspetto grammaticale nei bambini e la precocità della sensibilità alle strutture asettuali musicali. I bambini acquisiscono l'aspetto grammaticale in età molto precoce — prima dei due anni nelle lingue con aspetto obbligatorio come il russo — e la sensibilità alle strutture ritmiche e armoniche si sviluppa già nel primo anno di vita. C'è una connessione tra questi due sviluppi? I bambini che acquisiscono precocemente l'aspetto in una lingua con aspetto obbligatorio mostrano una maggiore sensibilità alle

strutture aspettuali musicali? Questa domanda è aperta e potenzialmente molto fruttuosa per la psicologia dello sviluppo.

La terza è l'estensione sistematica della teoria aspettuale alla dimensione elettroacustica e alla musica informatica. La musica elettroacustica — dalla musica su nastro delle origini al live electronics contemporaneo, dalla sintesi granulare alle composizioni assistite da algoritmi generativi — manipola il tempo interno degli eventi sonori a un livello di dettaglio e di controllo che nessuna tradizione strumentale può eguagliare: può costruire accomplishments di cui controlla la velocità di avvicinamento al terminus, può creare stati di durata arbitraria, può produrre achievement di dimensioni sub-percettive che non hanno equivalenti nella musica strumentale. Uno studio sistematico dell'aspetto nella musica elettroacustica richiederebbe strumenti di analisi e di formalizzazione più precisi di quelli sviluppati in questo saggio, e il suo risultato potrebbe retroilluminare la teoria aspettuale della musica strumentale mostrando in modo più esplicito strutture che nella musica strumentale emergono solo implicitamente.

Appendice A. Formalizzazione del Sistema Aspettuale Musicale

A.1 Presupposti e scelte notazionali

Le definizioni che seguono utilizzano una notazione in linguaggio pseudo-formale che è una semplificazione consapevole di sistemi formali più rigorosi — in particolare della semantica degli intervalli di Allen (1983) e del sistema di Krifka (1992) — ai fini dell'applicabilità analitica a casi musicali concreti. Questa scelta comporta la rinuncia a certi gradi di rigore tecnico (in particolare nella trattazione degli aspetti quantificazionali) in cambio di leggibilità e di applicabilità diretta. Chi desiderasse una formalizzazione più rigorosa può ricostruirla senza difficoltà a partire dalle definizioni informali presentate qui, usando i sistemi formali citati come base.

Sia E un evento musicale, $I(E) = [t_{\text{inizio}}, t_{\text{fine}}]$ il suo intervallo temporale di svolgimento, e $D(E) = t_{\text{fine}} - t_{\text{inizio}}$ la sua durata. Sia $TIPO(E, t)$ la funzione che assegna al tempo t la qualità aspettuale dell'evento E in quel momento. Sia $P = \{v_1, v_2, \dots, v_n\}$ un insieme di voci di un brano polifonico, e sia $A(v_i, t)$ la funzione che assegna a ciascuna voce v_i la sua Aktionsart al tempo t .

A.2 Definizione formale delle quattro categorie aspettuali

Uno **Stato** è un evento E tale che:

STATO(E) è vero se e solo se per ogni t nell'intervallo $I(E)$, il tipo dell'evento in quel momento è uguale al tipo dell'evento nel momento iniziale; non esiste un terminus interno all'evento; e la durata dell'evento non è strutturalmente determinata dall'interno ma dipende da condizioni esterne.

La condizione di omogeneità temporale cattura l'intuizione che lo stato "è già" ciò che è in ogni momento della sua durata: non ha una fase di sviluppo, non progredisce né si trasforma. La condizione dell'assenza di terminus interno cattura l'intuizione che uno stato non "tende" verso alcun punto di completamento. La condizione della durata non strutturalmente determinata cattura l'intuizione che uno stato può durare per qualsiasi intervallo di tempo senza che la sua struttura cambi tipo.

Un'**Attività** è un evento E tale che:

ATTIVITA(E) è vero se e solo se l'evento è dinamico (il tipo varia nel corso del suo svolgimento); per ogni sotto-intervallo dell'evento, il sotto-evento corrispondente è dello stesso tipo dell'evento intero; e non esiste un terminus interno all'evento.

La condizione di dinamicità distingue l'attività dallo stato. La condizione di subomogeneità — che qualsiasi parte di un'attività sia dello stesso tipo dell'intero — è la proprietà di cumulatività di Krifka: qualsiasi segmento di una corsa è ancora una corsa, qualsiasi concatenazione di corse è ancora una corsa. La condizione dell'assenza di terminus interno è condivisa con lo stato.

Un **Accomplishment** è un evento E tale che:

ACCOMPLISHMENT(E) è vero se e solo se l'evento è dinamico; esiste un terminus interno t_{term} che fa parte della struttura dell'evento; l'evento è direzionato verso t_{term} in ogni momento del suo intervallo; e l'interruzione dell'evento in qualsiasi momento t precedente a t_{term} produce un evento incompleto — un evento che manca del suo elemento definitorio.

La condizione di direzionalità cattura l'intuizione che in un accomplishment ogni momento "punta" verso il terminus: la struttura dell'evento è orientata, non omogenea. La condizione di incompletezza dell'interruzione è la più importante: distingue l'accomplishment dall'attività in modo operazionalmente verificabile. Se si interrompe un accomplishment prima del terminus, ciò che si ottiene non è un accomplishment più breve ma un evento di tipo diverso — categorialmente incompleto, non semplicemente abbreviato.

Un **Achievement** è un evento E tale che:

ACHIEVEMENT(E) è vero se e solo se esiste un istante t_{punto} tale che un cambiamento di stato si verifica a t_{punto} ; l'evento non ha parti rilevanti precedenti a

t_punto che appartengano alla sua struttura come achievement (le fasi preparatorie appartengono all'accomplishment precedente, non all'achievement); e lo stato delle cose dopo t_punto è categorialmente diverso dallo stato delle cose prima di t_punto in modo irreversibile.

La condizione di puntualità strutturale distingue l'achievement dall'accomplishment: non c'è un processo di raggiungimento ma un punto di cambiamento. La condizione di irreversibilità cattura la qualità cognitiva dell'achievement come momento che "cambia tutto": dopo l'achievement, il sistema è in uno stato diverso da prima, e questo cambiamento non può essere annullato (nella logica dell'evento, non necessariamente nella realtà fisica).

A.3 Telicità e atelicità come proprietà derivate

La telicità non è una quinta categoria aspettuale ma una proprietà che distingue gli accomplishments e gli achievement (telici) dagli stati e dalle attività (atelici):

TELICO(E) è vero se e solo se ACCOMPLISHMENT(E) oppure ACHIEVEMENT(E). Questo equivale a dire che esiste un terminus interno all'evento e che l'interruzione prima del terminus produce un evento incompleto.

ATELICO(E) è vero se e solo se STATO(E) oppure ATTIVITA(E). Questo equivale a dire che non esiste un terminus interno all'evento, oppure che l'eventuale terminus è estrinseco — determinato da condizioni esterne alla struttura dell'evento — e non strutturale.

In termini operazionali, la distinzione può essere catturata attraverso il *test di cumulatività*: un evento è atelico se la somma di due sue istanze consecutive è ancora un'istanza dello stesso tipo di evento. Un pedale di tonica più un secondo pedale di tonica è ancora un pedale di tonica (più lungo). Una cadenza autentica più una seconda cadenza autentica consecutiva sono due cadenze distinte, non una cadenza più lunga.

Una precisazione formalmente importante: la telicità di un evento è relativa al livello di descrizione scelto. Come si è mostrato nell'analisi del Preludio BWV 846 di Bach, un evento può essere atelico al livello della struttura di superficie e telico al livello della struttura armonica profonda. La formalizzazione deve dunque indicare esplicitamente a quale livello di descrizione si applica: TELICO(E, livello) dove *livello* è un parametro che specifica il livello di analisi (superficie, struttura armonica locale, struttura tonale profonda, struttura formale di macro-livello).

A.4 La diatesi aspettuale: formalizzazione

Una **configurazione di diatesi** a un tempo t è la mappa che assegna a ciascuna voce v_i del brano la sua Aktionsart al tempo t . Sia DIATESI(t) questa mappa.

La **sincronia aspettuale** al tempo t si ha quando tutte le voci del brano si trovano nella stessa Aktionsart al tempo t : per ogni coppia di voci v_i e v_j , l'Aktionsart di v_i a t è uguale all'Aktionsart di v_j a t .

La **dissincronia aspettuale** al tempo t si ha quando la condizione di sincronia non è soddisfatta.

Il **grado di dissincronia** al tempo t è il numero di Aktionsart distinte presenti tra le voci del brano al tempo t , normalizzato rispetto al numero totale di voci. Il grado minimo (uguale a uno diviso il numero di voci) si ha quando tutte le voci sono nella stessa Aktionsart; il grado massimo (uguale a uno) si ha quando ogni voce è in un'Aktionsart diversa da tutte le altre.

La **tensione aspettuale** al tempo t è una funzione del grado di dissincronia e della distanza media al terminus per le voci che si trovano in accomplishment al tempo t . La tensione aspettuale è massima quando il grado di dissincronia è alto — molte voci in Aktionsart diverse — e le voci in accomplishment sono lontane dai loro terminus. La tensione aspettuale è minima (ma non necessariamente nulla) quando tutte le voci sono nella stessa Aktionsart: una sincronia di tutti in stato produce una tensione aspettuale molto bassa; una sincronia di tutti in accomplishment produce una tensione aspettuale proporzionale alla distanza dal terminus condiviso.

A.5 La coercion aspettuale: formalizzazione dei tre tipi fondamentali

La coercion aspettuale musicale è la trasformazione dell'Aktionsart percepita di un evento in risposta al contesto in cui è collocato. Tre tipi di coercion sono particolarmente rilevanti per l'analisi musicale.

La **coercion telicizzante** si verifica quando un'attività intrinsecamente atelica è collocata in un contesto di forte aspettativa cadenzale, che la forza a essere percepita come un accomplishment in corso. La struttura sonora dell'evento non cambia, ma la sua Aktionsart cambia perché il contesto formale impone una diversa struttura temporale. Il caso tipico è l'ostinato di basso sullo sviluppo sonatistico, che viene percepito come direzionato verso il terminus della ripresa non per via della sua struttura intrinseca ma per via della pressione teleologica esercitata dal contesto formale.

La **coercion atelicizzante** si verifica quando un accomplishment viene ripetuto un numero di volte sufficiente a trasformarlo in un'attività iterativa o abituale. Il caso tipico è la ripetizione multipla della cadenza finale nelle conclusioni delle sinfonie classiche: la prima cadenza è percepita come accomplishment che raggiunge il terminus; le ripetizioni successive trasformano retroattivamente la prima in parte di un pattern iterativo, e da quel momento le cadenze non sono più percepite come terminus ma come unità di un processo dichiarativo atelico.

La **stativizzazione** si verifica quando un accomplishment il cui terminus viene continuamente differito — attraverso l'estensione del percorso verso il terminus al di là della soglia di aspettativa dell'ascoltatore — viene percepito come uno stato di tensione irrisolta. Il caso tipico è il lungo pedale di dominante che precede la ripresa in molti sviluppi sonatistici: la dominante è strutturalmente un accomplishment in corso (è orientata verso la tonica), ma la sua prolungazione trasforma la percezione da "accomplishment vicino al termine" a "stato di tensione che vale", in un processo di stativizzazione graduale che Beethoven gestisce con coscienza particolarmente elaborata.

A.6 Protocollo di analisi aspettuale di un brano polifonico

Il seguente protocollo in sei fasi offre una guida sistematica per l'analisi aspettuale di un brano contrappuntistico.

Nella **prima fase** si procede alla segmentazione: il brano viene diviso in unità di analisi — frasi, periodi, sezioni — e per ciascuna unità si identifica il livello di analisi aspettuale rilevante: micro (singola frase), meso (sezione), macro (forma intera). È importante scegliere il livello di analisi in modo consapevole e dichiararlo esplicitamente, poiché la stessa struttura può avere Aktionsart diverse a livelli diversi.

Nella **seconda fase** si classificano gli eventi locali: per ciascuna voce e per ciascuna unità di analisi, si applicano i quattro criteri operazionali del paragrafo 2.1 — sospensibilità senza incoerenza, cumulatività, direzionalità interna, irriducibilità dell'endpoint — per classificare l'Aktionsart dell'evento. I criteri devono essere applicati in modo coerente e i casi ambigui devono essere esplicitamente discussi.

Nella **terza fase** si mappa la configurazione di diatesi: per ciascun istante strutturalmente rilevante — inizio e fine di ciascuna unità, momenti cadenzali, entrate di soggetti in fuga — si costruisce la configurazione di diatesi e si calcola il grado di dissincronia.

Nella **quarta fase** si calcola la curva di tensione aspettuale per una serie di istanti nel corso del brano, producendo una curva che rappresenta l'andamento della dissincronia nel tempo. Questa curva viene poi confrontata con la curva di tensione armonica (calcolabile attraverso il tonal pitch space di Lerdahl) e con la curva di tensione ritmica (calcolabile attraverso la teoria del metro di Hasty), per verificare le somiglianze e le differenze tra le tre dimensioni della tensione musicale.

Nella **quinta fase** si identificano i fenomeni di coercion: si cercano strutture ateliche in contesti di forte aspettativa cadenzale (coercion telicizzante); si cercano accomplishments ripetuti che acquistano struttura iterativa (coercion atelicizzante); si cercano accomplishments bloccati che si stativizzano (stativizzazione). Per ciascun caso di coercion identificato, si descrive il meccanismo contestuale che la produce.

Nella **sesta fase** si sintetizza la narrativa aspettuale del brano: come cambia la configurazione di diatesi nel corso del brano, quali sono i momenti di massima sincronia e dissincronia, come si relaziona la narrativa aspettuale con la narrativa armonica e formale. La narrativa aspettuale non sostituisce né supera la narrativa armonica: la integra, aggiungendo una dimensione cognitiva che la sola analisi armonica non può catturare.

Appendice B. Il Rapporto con la Teoria di Hasty: Convergenze e Divergenze

La teoria di Hasty in *Meter as Rhythm* (1997) e la teoria aspettuale sviluppata in questo saggio convergono su tre punti fondamentali e divergono su due.

Le convergenze riguardano la struttura profonda del tempo musicale. La prima convergenza è l'idea che il tempo musicale sia strutturato dall'interno degli eventi, non imposto da una griglia esterna: per Hasty, il metro emerge dalla proiezione; per questa monografia, l'aspetto è una proprietà interna degli eventi che non può essere ridotta a proprietà esterne come la posizione nella forma o il tipo di accordo. La seconda convergenza è l'idea che l'inizio di un evento sia un atto di apertura di una durata prospettica: il concetto hastyiano di "becoming" — ogni evento è sempre in processo di diventare, di aprire una durata futura — corrisponde esattamente alla struttura degli accomplishments vendleriani, che "aprono" il loro processo verso il terminus nel momento stesso in cui iniziano. La terza convergenza è l'idea che la tensione musicale sia la tensione tra la proiezione attesa e l'evento reale: la "incompleteness" hastyiana — la sensazione che un evento metrico sia aperto, in attesa del suo completamento — corrisponde alla tensione aspettuale della diatesi dissincronica.

Le divergenze riguardano il dominio di applicazione e il problema della polifonia. La prima divergenza è che Hasty analizza primariamente il livello ritmico-metrico, mentre questa monografia analizza primariamente il livello armonico-funzionale: le due analisi operano su livelli diversi che interagiscono ma non si sovrappongono interamente. Non è chiaro, nella teoria di Hasty, in che modo la struttura metrica e la struttura armonica si relazionino e in che modo i loro rispettivi regimi aspettuali interagiscano quando sono in conflitto o in accordo. Questa relazione è uno dei problemi aperti più produttivi che l'integrazione dei due framework può affrontare.

La seconda divergenza è la più fondamentale: Hasty non tematizza la polifonia come problema di configurazioni di diatesi simultanee. La sua teoria è fondamentalmente monodica nella prospettiva, anche quando è applicata a strutture polifoniche: analizza come un singolo filo temporale — il filo del metro, della proiezione — si struttura

aspettualmente, non come fili temporali multipli e simultanei interagiscono nelle loro diverse strutture aspettuali. La diatesi aspettuale è dunque un'estensione della teoria di Hasty in una direzione che Hasty stesso non ha sviluppato, non una sua conseguenza diretta. Integrare il concetto di proiezione hastyana con il concetto di diatesi multipla sviluppato qui è uno dei compiti teorici più ambiziosi che la ricerca futura potrebbe intraprendere.

Appendice C. La Stratificazione Aspettuale nella Musica Vocale

Questa appendice approfondisce il tema trattato nel paragrafo 2.6 del Capitolo II, sviluppando in modo più sistematico i meccanismi attraverso cui i due livelli aspettuali della musica vocale — l'aspetto grammaticale del testo e l'aspetto musicale della composizione — interagiscono, si sovrappongono e talvolta confliggono.

La sincronizzazione aspettuale può avvenire a tre livelli di granularità: a livello di sillaba (ogni sillaba carica di significato aspettuale riceve un trattamento musicale aspettualmente corrispondente), a livello di parola (ogni parola verbale riceve un trattamento musicale corrispondente alla sua Aktionsart lessicale), o a livello di frase (l'intera struttura retorica della frase riceve un trattamento musicale corrispondente alla sua Aktionsart complessiva). Il madrigale rinascimentale pratica prevalentemente la sincronizzazione a livello di sillaba e di parola; il corale luterano bachiano pratica prevalentemente la sincronizzazione a livello di frase; il recitativo barocco pratica la sincronizzazione a livello di parola per i verbi principali e a livello di frase per le strutture sintattiche.

Il controaspetto compositivo — la sincronizzazione volutamente errata tra aspetto del testo e aspetto della musica — produce effetti espressivi che dipendono dalla qualità del conflitto. Quattro tipi di conflitto sono analiticamente distinguibili. Il primo tipo è il conflitto di telicità: un testo con verbi telici (passato remoto, verbi di completamento) viene intonato su musica atelica (armonie prolungate, linee melodiche non cadenzanti), come nell'aria mozartiana analizzata nel Capitolo II. Il secondo tipo è il conflitto di aspetto esterno: un testo con verbi perfettivi (evento visto dall'esterno nella sua interezza) viene intonato su musica imperfettiva (evento presentato dall'interno del suo svolgimento), come in certi recitativi tardobarocchi dove l'accompagnamento tende a "stare dentro" l'evento anche quando il testo lo presenta come concluso. Il terzo tipo è il conflitto di Aktionsart: un testo con verbi stativi viene intonato su musica dinamico-telica, producendo l'effetto di una calma di facciata che la musica contraddice con la sua direzionalità, come in certi lieder schubertiani in cui la

scena descritta nel testo è pacifica ma l'accompagnamento pianistico è inquieto. Il quarto tipo è il conflitto di durata: un testo che descrive eventi rapidi e puntuali viene intonato su musica di grande sviluppo temporale, o viceversa.

L'analisi del conflitto aspettuale nella musica vocale ha implicazioni teoriche che vanno al di là dell'analisi dei singoli brani. Essa mostra che l'aspetto grammaticale del testo e l'aspetto musicale della composizione non sono due sistemi autonomi che si sovrappongono fortuitamente: sono due sistemi di organizzazione del tempo interno degli eventi che condividono le stesse categorie cognitive fondamentali — telico/atelico, perfettivo/imperfettivo, stativo/dinamico, puntuale/durativo — realizzandole in media diversi. La loro co-presenza in un brano vocale è dunque qualcosa di più di una semplice sovrapposizione: è la confluenza di due manifestazioni dello stesso sistema cognitivo in un unico spazio espressivo, e la loro relazione — sincronia, dissincronia, conflitto, risoluzione — è essa stessa una dimensione del significato musicale-vocale.

Tavola Tipologica delle Tradizioni Musicali

La tavola che segue sintetizza la tipologia aspettuale sviluppata nel Capitolo III, con l'aggiunta dei dati comparativi provenienti dal Capitolo V. Le colonne indicano il tipo di Aktionsart privilegiata come nucleo strutturale; il grado di grammaticalizzazione (da bassa, che indica preferenza stilistica, ad alta, che indica obbligatorietà sistemica); il sito strutturale della grammaticalizzazione; e il tipo prevalente di configurazione di diatesi.

Il **contrappunto palestriniano** privilegia un'atelicità di fondo con telicità locale nelle clausole cadenzali; il grado di grammaticalizzazione è medio, poiché l'aspetto è marcato nei luoghi strutturali chiave ma lasciato non marcato nel tessuto interstitiale; il sito strutturale sono le clausole cadenzali su finalis, confinalis e terza del modo; la configurazione di diatesi prevalente è mista e fluttuante, con alternanza frequente di stati e attività nelle diverse voci.

Il **contrappunto bachiano** privilegia la telicità a tutti i livelli strutturali; il grado di grammaticalizzazione è alto, con funzionalità armonica obbligatoria in ogni momento; il sito strutturale è la gerarchia cadenzale integrata con la struttura tonale; la configurazione di diatesi prevalente è una progressione verso la sincronia crescente con l'avvicinarsi delle cadenze.

Il **classicismo viennese** privilegia la telicità graduata gerarchicamente; il grado di grammaticalizzazione è molto alto, con la struttura PAC/IAC/HC obbligatoriamente

determinata dalla posizione formale; il sito strutturale è la gerarchia cadenzale nella forma sonatistica; la configurazione di diatesi prevalente è la sequenza stato-attività-accomplishment-achievement nella frase.

Il **romanticismo tardo** privilegia una telicità sospesa, con termine continuamente differito attraverso il cromatismo e le risoluzioni evase; il grado di grammaticalizzazione è medio-alto, poiché la telicità è ancora presente come norma di fondo ma è sistematicamente ostacolata; il sito strutturale è la macro-forma con micro-eventi spesso atelici o stativi; la configurazione di diatesi prevalente è alta dissincronia locale con convergenza teleologica di grande scala.

Il **minimalismo americano** privilegia l'atelicità come struttura organizzativa dominante; il grado di grammaticalizzazione è alto, poiché la struttura iterativa obbligatoria esclude i terminus come categorie organizzative; il sito strutturale è la struttura iterativa dei moduli; la configurazione di diatesi prevalente è la sincronia atelica uniforme tra le voci.

Il **jazz bebop** privilegia l'accomplishment rapido e denso; il grado di grammaticalizzazione è alto, con la progressione accordale obbligatoria che determina la struttura telica di ogni frase; il sito strutturale è la progressione accordale con le sue funzioni ii-V-I; la configurazione di diatesi prevalente è la sincronia di accomplishments densi e ravvicinati.

Il **jazz modale** privilegia un'atelicità di fondo con accomplishments sospesi e tenuti deliberatamente in stato di indeterminazione; il grado di grammaticalizzazione è medio, poiché la scala modale come sfondo è obbligatoria ma la struttura aspettuale delle frasi è lasciata all'improvvisatore; la configurazione di diatesi prevalente è la dissincronia deliberata con rifiuto sistematico della risoluzione.

Il **tala indostanico** privilegia la telicità ciclica con terminus sul sam; il grado di grammaticalizzazione è molto alto, poiché la struttura del tala determina obbligatoriamente il terminus di ogni ciclo; il sito strutturale è il ciclo del tala con il sam come terminus intrinseco; la configurazione di diatesi prevalente è l'accomplishment ciclico iterato.

Il **gamelan giavanese** privilegia la stratificazione aspettuale per scale temporali diverse; il grado di grammaticalizzazione è alto, poiché ogni strumento ha un livello temporale obbligatoriamente determinato dalla sua posizione nella gerarchia colotomia; il sito strutturale è la struttura colotomica con i gong come marcatori di livello; la configurazione di diatesi prevalente è la pluralità aspettuale stratificata per scala temporale.

Le **polifonie di interlocking africane** privilegiano la telicità distribuita tra le voci dell'ensemble; il grado di grammaticalizzazione è alto, poiché ogni parte è strutturalmente incompleta senza le parti complementari; il sito strutturale è il ciclo dell'ensemble come unità, non le singole parti; la configurazione di diatesi prevalente è

la convergenza obbligatoria di frammenti aspettualmente distribuiti verso un terminus collettivo.

Integrazioni alla monografia

Quattro zone di approfondimento

I. Gli stati musicali: una difesa più elaborata

La categoria più controversa

Tra le quattro Aktionsart che questo saggio ha applicato alla musica, gli stati sono quella che incontra la resistenza più istintiva e più difficile da smontare. L'obiezione di fondo non è tecnica ma quasi fenomenologica: la musica è movimento, è tempo che scorre, è cambiamento. L'idea che in musica possa esserci qualcosa di analogo allo "stato" — una condizione che semplicemente "vale" senza progredire né tendere — sembra contraddire la natura stessa del mezzo. Anche una nota lunga tenuta si svolge nel tempo, occupa durata, ha un attacco e un rilascio: come può essere "statica" in qualsiasi senso non banale?

L'obiezione è seria e merita una risposta articolata a tre livelli distinti, che riguardano rispettivamente la definizione operativa degli stati, la loro realizzazione concreta nella musica, e il loro ruolo cognitivo nell'esperienza dell'ascolto.

La statività come qualità funzionale, non fisica

Il primo livello della risposta riguarda la definizione. Quando si dice che uno stato musicale è atelico e omogeneo, non si sta descrivendo una proprietà fisica del suono — l'assenza di variazioni nel segnale acustico, che sarebbe effettivamente impossibile in qualsiasi musica reale — ma una proprietà della struttura funzionale dell'evento: il fatto che in ogni momento della sua durata l'evento "è già" ciò che è, che non ha fasi di sviluppo verso qualcosa di diverso, che il suo valore strutturale non dipende da ciò che verrà dopo.

Questa distinzione tra proprietà fisica e proprietà funzionale non è un'invenzione ad hoc: è precisamente la distinzione che la linguistica fa tra la struttura superficiale di un

enunciato e la sua struttura semantica profonda. Il verbo italiano "sapere" in "so suonare il pianoforte" descrive uno stato anche se il parlante, mentre lo pronuncia, sta fisicamente svolgendo l'atto dinamico del parlare. La statività non è una proprietà del segnale fisico ma della struttura semantica: "sapere" descrive una condizione che vale in modo omogeneo nel corso della sua durata, senza fasi di sviluppo verso un terminus. Analogamente, un pedale di tonica nella musica di Bach è uno stato funzionale anche se il suono fisicamente varia, decade, viene reattaccato: ciò che conta è che la funzione armonica che quella nota realizza — la stabilizzazione del centro tonale — è già pienamente realizzata nel primo momento e rimane invariata per tutta la durata del pedale.

Cinque realizzazioni concrete degli stati musicali

Il secondo livello della risposta consiste nel mostrare che gli stati musicali non sono un'astrazione teorica ma una categoria che cattura strutture concrete e diversificate della pratica compositiva.

La prima realizzazione è il **pedale di tonica**, già discusso. Il basso che rimane fermo sulla nota fondamentale del modo, mentre le voci superiori possono ornamentare, armonizzare, anche dissonare brevemente, realizza uno stato funzionale di stabilità tonale. La sua qualità caratteristica è precisamente quella dell'omogeneità: ogni momento del pedale di tonica "vale" quanto il precedente, e il primo momento e l'ultimo sono strutturalmente equivalenti. Nessuna parte del pedale è "più vicina al terminus" di un'altra perché non c'è terminus: il pedale finisce quando finisce, per ragioni esterne alla sua struttura.

La seconda realizzazione è la **zona di liquidazione**, nella terminologia di Schoenberg. Nelle composizioni classiche, tra due sezioni principali si trova spesso una zona di transizione in cui il materiale tematico si dissolve — le figure si fanno più brevi, più semplici, meno direzionali — producendo una condizione di attesa che non è orientata verso nessun terminus specifico ma che "è" semplicemente un momento di sospensione funzionale. Questa zona di liquidazione ha la struttura di uno stato: la sua funzione è di "stare" tra una sezione e l'altra senza tendere attivamente verso la successiva.

La terza realizzazione è il **drone** delle tradizioni modali e di molte musiche folk e non-occidentali: la nota tenuta che funge da sfondo permanente su cui si svolge tutta la melodia. Il drone è lo stato musicale per eccellenza, e la sua universalità crossculturale — si trova nella cornamusa scozzese, nel tambura indiano, nella vihuela del flamenco, nel bordone gregoriano — suggerisce che la categoria cognitiva dello stato musicale è genuinamente universale, non specificamente legata alla tradizione tonale occidentale.

La quarta realizzazione è la **ripetizione ostinata di figura** nel minimalismo: il loop di Philip Glass, l'ostinato di piano di *Metamorphosis* ripetuto per decine di minuti, ha la

struttura funzionale di uno stato anche se la struttura fisica è dinamica. La ripetizione rimuove la direzionalità: dopo le prime poche iterazioni, l'ascoltatore cessa di percepire il pattern come un processo che "sta andando da qualche parte" e comincia a percepirlo come una condizione che "vale". La transizione da attività atelica a stato è graduale — la coercion stativizzante agisce nel tempo — ma il risultato finale è percettivamente distinto.

La quinta realizzazione è meno ovvia ma forse la più importante teoricamente: la **zona post-cadenziale**. Subito dopo il completamento di un accomplishment cadenzale, c'è quasi sempre — anche nella musica più dinamica — un brevissimo momento in cui l'accordo di tonica "risuona" senza tendere ancora verso nessun altro terminus, in cui la struttura funzionale è quella di un post-achievement stativo: si è arrivati, e per un momento si "è" semplicemente nel punto di arrivo. Questo stato post-cadenziale è brevissimo — può durare meno di una misura — ma è reale e cognitivamente distinto dal processo che lo ha preceduto. È il respiro dopo la risoluzione, la pausa che separa un accomplishment dall'inizio del successivo.

Il ruolo cognitivo degli stati nell'esperienza dell'ascolto

Il terzo livello della risposta riguarda la funzione cognitiva degli stati nell'esperienza dell'ascolto. Se gli stati non esistessero come categoria distinta, ogni momento della musica sarebbe percepito come parte di un processo direzionato verso qualcosa. Ma l'esperienza effettiva dell'ascolto non è questa: ci sono momenti in cui si "è" nella musica senza aspettare nulla di specifico, momenti di stabilità non come assenza di tensione ma come tipo qualitativamente diverso di presenza temporale. Questi momenti — il pedale di tonica prolungato prima di una fuga, il drone dell'introduzione di una raga, la ripetizione estatica del minimalismo — producono uno stato cognitivo che molti ascoltatori descrivono come "sospensione del tempo", "assenza di aspettativa", "presenza pura". Questa fenomenologia non è casuale né puramente soggettiva: è la risposta cognitiva alla struttura aspettuale stativa dell'evento musicale, alla sua qualità di "valere" piuttosto che "tendere".

La distinzione cognitiva tra stati e processi musicali è verificabile sperimentalmente attraverso misurazioni della risposta di orientamento auditiva (*auditory orienting response*): si è mostrato che la risposta di orientamento — il correlato neurale della sorpresa e dell'aspettativa infranta — è significativamente diversa quando viene interrotta una struttura stativa rispetto a quando viene interrotta una struttura processuale. L'interruzione di un pedale di tonica non produce la stessa risposta cognitiva dell'interruzione di un accomplishment cadenzale a metà strada: la prima è una sorpresa di tipo "qualcosa ha smesso di essere quello che era", la seconda è una sorpresa di tipo "il processo è stato tagliato prima di completarsi". Questa differenza qualitativa nella risposta cognitiva all'interruzione è la prova empirica più diretta che stati e processi sono categorie cognitive distinte nell'ascolto musicale.

II. La musica post-tonale: quando la grammatica della telicità viene abolita

Il problema

La domanda che il serialismo e la musica post-tonale pongono alla teoria asettuale è la più radicale che il framework debba affrontare: se la telicità armonica nel sistema tonale è realizzata attraverso la progressione verso la tonica — se gli accomplishments musicali "puntano" verso la tonica come terminus intrinseco — che cosa succede dell'aspetto quando la tonica viene abolita? Il sistema dodecafonico di Schoenberg, il serialismo integrale di Boulez, la musica per cluster di Ligeti, il teatro strumentale di Nono: questi sistemi eliminano deliberatamente la struttura tonale gerarchica che era il veicolo principale della telicità nella musica occidentale colta. La teoria asettuale sopravvive a questa abolizione?

La risposta proposta qui è che la teoria sopravvive, ma richiede di spostare il sito della telicità da quello armonico ad altri domini strutturali che ogni tradizione post-tonale ha sviluppato come sostituti funzionali della cadenza. Questo spostamento non è un ripiego: rivela che la telicità non era mai strettamente dipendente dall'armonia tonale, ma era realizzata attraverso l'armonia tonale come uno dei suoi veicoli possibili. La struttura cognitiva fondamentale — processo con terminus intrinseco — è più profonda della sua realizzazione armonica, e può essere realizzata attraverso la dinamica, il registro, la densità, la dissoluzione timbrica, la forma spaziale del suono.

Schoenberg: la telicità trasferita alla forma motivica

Nelle opere dodecafoniche di Schoenberg — dalla *Suite per pianoforte* op. 25 alla *Terza sonata* (incompiuta) al *Concerto per violino* op. 36 — la telicità non scompare ma si trasferisce dalla dimensione armonica alla dimensione motivica e formale. La serie dei dodici suoni, in quanto struttura chiusa e completa in sé stessa, non produce la direzionalità cumulativa della progressione cadenzale; ma le trasformazioni della serie — inversione, retrogradazione, inversione del retrogrado, trasposizioni — producono processi che hanno la struttura degli accomplishments: partono da una configurazione iniziale, si sviluppano attraverso trasformazioni sistematiche, e "arrivano" a configurazioni di riferimento che fungono da terminus strutturali.

Nelle variazioni per pianoforte op. 31, ogni variazione è un accomplishment nel senso formale: parte dalla configurazione iniziale della serie, la elabora secondo procedure definite, e raggiunge una configurazione di arrivo che marca il completamento della variazione. Il terminus non è tonale — non c'è una tonica verso cui tendere — ma è motivico-formale: è il momento in cui la trasformazione della serie si è completata

secondo la logica interna della variazione. La direzionalità è prodotta non dall'armonia funzionale ma dalla logica della trasformazione seriale, che ha la stessa struttura cognitiva dell'accomplishment pur realizzandosi in un dominio completamente diverso.

La conquista teorica di questo riconoscimento è significativa: la telicità è una struttura cognitiva che può essere realizzata in domini diversi dall'armonia funzionale. Schoenberg non ha abolito la telicità; ha trovato un nuovo veicolo per essa nel sistema seriale.

Webern: la telicità nel gesto puntuale

Anton Webern rappresenta il caso estremo nell'economia aspettuale del serialismo. Le sue opere più mature — le *Variazioni per pianoforte* op. 27, le *Cantate* op. 29 e 31, il *Concerto* op. 24 — hanno una densità di eventi così alta e una durata degli eventi così breve che la struttura aspettuale dominante non è né l'accomplishment né l'attività ma l'achievement: quasi ogni evento sonoro è puntuale, isolato, privo di fase preparatoria visibile, circondato da silenzio. Il suono weberniano è quasi sempre un achievement: un cambiamento di stato che avviene senza processo precedente percepibile, un punto nel silenzio.

Questa predilezione per l'achievement ha conseguenze sulla struttura della tensione. Un brano costruito quasi esclusivamente su achievements — come i *Bagatelles* op. 9 per quartetto d'archi — produce una tensione aspettuale di tipo particolare: non la tensione dell'accomplishment che "costruisce verso", che accumula direzionalità nel corso del suo svolgimento, ma la tensione del silenzio che precede l'achievement ignoto, la tensione di non sapere quando, dove e di quale tipo sarà il prossimo cambiamento di stato. Questa è la "Angst" weberniana in termini aspettuati: non la tensione del processo incompiuto ma la tensione dell'attesa senza struttura di accomplishment visibile.

Il trattamento del silenzio in Webern è, da questo punto di vista, il trattamento dello stato puro: i silenzi che separano gli eventi sonori non sono pause nel processo (come le pause nelle progressioni tonali, che sono parte della struttura dell'accomplishment in corso) ma stati veri e propri, momenti in cui la musica "è" nel silenzio senza tendere verso nessun terminus specifico. La diatesi aspettuale di Webern è dunque una costante alternanza di stati (silenzi) e achievements (suoni puntuali), con quasi nessun accomplishment nel senso pieno.

Boulez e il serialismo integrale: la telicità nella struttura parametrica

Il serialismo integrale di Pierre Boulez — nelle *Structures Ia* per due pianoforti (1952), nel primo libro delle *Sonates* per pianoforte, e più compiutamente in *Le Marteau sans maître* (1955) — pone alla teoria aspettuale un problema diverso da Schoenberg e

Webern. Nel serialismo integrale, non solo le altezze ma anche le durate, le intensità e i modi di attacco sono serializzati: ogni parametro del suono è sottoposto a un ordinamento seriale che determina in modo rigido la successione dei suoi valori. Qual è la struttura aspettuale di un sistema in cui ogni dimensione del suono è governata da una serie?

La risposta dipende dal livello di analisi. A livello micro — il livello del singolo evento sonoro — il serialismo integrale tende a produrre una proliferazione di achievements: eventi puntuali e isolati in cui parametri multipli cambiano simultaneamente, producendo cambiamenti di stato su più dimensioni contemporaneamente. A livello meso — il livello della frase o della sezione — il risultato è spesso una texture di alta densità in cui la struttura aspettuale è difficile da identificare: le serie di durata, altezza e intensità si sovrappongono in modo tale che non è chiaro in quale direzione nessuno dei parametri stia "tendendo". Questo produce uno dei regimi aspettuati più peculiari della musica del Novecento: una condizione di alta dinamicità e alta densità di cambiamento in cui tuttavia non si percepisce alcuna direzionalità complessiva verso un terminus. In termini aspettuati, il serialismo integrale nel suo stadio più rigoroso produce una forma di atelicità dinamica: è un'attività nel senso di Vendler, ma un'attività con densità di cambiamento molto più alta di qualsiasi attività nel repertorio tonale.

Boulez stesso non era soddisfatto di questo risultato, e la sua produzione successiva si allontana progressivamente dalla rigidità del serialismo integrale verso strutture più flessibili che permettono la creazione di archi di tensione a media scala. *Pli selon pli* (1957-1962) e *Répons* (1981-1984) introducono strutture formali che producono accomplishments di macro-livello — processi con terminus formali riconoscibili — anche se la loro superficie seriale conserva la alta densità di achievements della prima maniera. Questo movimento dalla atelicità dinamica al recupero degli accomplishments di macro-livello è la traiettoria compositiva della maturità bouleziana vista attraverso la lente aspettuale.

Ligeti: l'aspetto nella micropolifonia e nel flusso temporale

György Ligeti ha sviluppato un linguaggio compositivo che è, tra tutti quelli del Novecento, quello in cui le categorie aspettuati si manifestano più chiaramente come principi organizzativi consapevoli. La micropolifonia di *Atmosphères* (1961) e *Lontano* (1967) produce texture in cui un numero enorme di voci si muovono in modo indipendente, con ritmi leggermente diversi, creando un flusso sonoro complessivo che non può essere analizzato in termini di eventi discreti ma che ha una struttura aspettuale globale chiaramente identificabile: è un'attività di macro-livello — un processo omogeneo e cumulativo senza terminus intrinseco — costruita dall'intreccio di accomplishments e attività individuali così minutamente sfasati da perdersi nella texture complessiva.

Nelle opere per pianoforte degli *Études* (1985-2001), Ligeti esplora la diatesi aspettUALE nella polifonia con una consapevolezza quasi teorica. Il *Désordre* (Étude n. 1) sovrappone due mani che suonano lo stesso pattern ritmico-melodico con accenti metrici diversi e sfasamenti progressivi, producendo una diatesi di due accomplishments paralleli che si de-sincronizzano progressivamente. Aspettualmente, è come se due voci che partono dalla stessa configurazione aspettUALE — entrambe in accomplishment, entrambe orientate verso il medesimo terminus — perdessero gradualmente la sincronia dei rispettivi termini, creando una dissonanza aspettUALE crescente che è indipendente dalla dissonanza armonica. *L'escalier du diable* (Étude n. 13) produce invece un accomplishment di scala sinfonica — un processo ascendente e accelerante verso un terminus implicito che non arriva mai, che supera il registro dello strumento e continua "oltre" in un'estensione immaginaria — che è forse il caso più estremo di accomplishment senza achievement terminale nella letteratura pianistica.

Nono: l'aspetto come dimensione politica del tempo

Luigi Nono occupa un posto speciale in questa analisi per ragioni che vanno al di là della teoria musicale. La sua riflessione sul tempo musicale — espressa sia nella musica che negli scritti teorici, in particolare nei testi raccolti in *Verso Prometeo* (1984) e negli ultimi quaderni — è una riflessione esplicitamente aspettUALE, anche se Nono non usa mai questo termine. Il suo concetto di "tempo interiore del suono" — il modo in cui ogni suono ha una propria struttura temporale interna che la notazione tradizionale non può catturare — è precisamente la nozione di Aktionsart applicata al suono elettronico e al suono strumentale amplificato.

Nelle opere tardive — *Prometeo* (1984), il *Diario polacco n. 2* (1982), le *Corde di dialogo disteso* (1989) — Nono costruisce strutture aspettUALI radicalmente diverse dall'economia classica. Il suono elettronico in *Prometeo* — manipolato in tempo reale dagli ingegneri del suono dell'IRCAM su indicazione del direttore — ha strutture aspettUALI che cambiano in tempo reale: un suono inizia come stato (ampio, stazionario, privo di direzionalità) e può essere trasformato, attraverso la manipolazione elettronica, in un accomplishment (acquista direzionalità, si orienta verso un punto di arrivo) o in un achievement (viene tagliato, cambia registro, cambia timbro bruscamente). Questa manipolazione in tempo reale dell'Aktionsart del suono — il fatto che la struttura aspettUALE di un evento possa essere modificata nel corso del suo svolgimento da un agente esterno — è una delle pratiche compositive più radicali del Novecento in termini di teoria del tempo.

III. La verifica empirica: analisi sistematica dell'Andante della Sonata K. 331

Perché la K. 331

La scelta del primo movimento della Sonata per pianoforte in la maggiore K. 331 di Mozart come caso di analisi sistematica seguendo il protocollo dell'Appendice A risponde a tre criteri. Il primo è la struttura del movimento — un tema con variazioni — che permette di analizzare come la stessa struttura melodica si trasformi aspettualmente attraverso le variazioni, offrendo una variabile controllata. Il secondo è la monofonia: è musica per pianoforte senza polifonia strumentale vera e propria, il che permette di concentrarsi sulla struttura aspettuale della singola linea armonica senza complicare l'analisi con la diatesi multi-voce. Il terzo è la sua centralità nel repertorio analitico — è stato analizzato da Schenkerian, da teorici della forma, da cognitivisti — e il confronto con le analisi esistenti può mostrare cosa la prospettiva aspettuale aggiunge o modifica.

Fase 1: segmentazione

Il movimento è diviso in tema più sei variazioni. Il tema ha 18 misure in forma binaria semplice (A: 8 misure, B: 10 misure). Ogni variazione conserva questa struttura binaria. Il livello di analisi scelto è **meso** (la frase) con attenzione al livello micro nelle misure strutturalmente rilevanti. Il livello macro (il movimento intero) verrà discusso nella fase 6.

Fase 2: classificazione degli eventi locali nel tema

Misure 1-4 (presentazione, sezione A): il tema si apre con un'armonia di tonica in la maggiore in posizione fondamentale, con una melodia che si muove per terza discendente e poi risale. Applicando i quattro criteri operazionali: l'evento è sospendibile senza incoerenza (si potrebbe fermare qui senza che la struttura diventi impossibile); è cumulativo (si può allungare questa fase senza che cambi tipo); manca di direzionalità interna (non sta costruendo verso un terminus specifico); l'endpoint è irrilevante (la misura 4 non realizza un terminus strutturale ma è solo una tappa). **Aktionsart: stato.**

Misure 5-8 (continuazione, sezione A): l'armonia inizia a muoversi più rapidamente, con un movimento ii-V-I alla misura 8 che realizza la cadenza autentica perfetta. Le misure 5-7 sono un'attività (movimento armonico senza terminus ancora chiaro), le misure 7.5-8 sono il passaggio da attività ad accomplishment in chiusura,

con il terminus sulla cadenza PAC. **Aktionsart: attività in transizione ad accomplishment.**

Misure 9-12 (inizio sezione B): il tema si sposta sulla dominante mi maggiore. Aspettualmente, questa sezione ha la struttura di un nuovo accomplishment che si apre: l'arrivo sulla dominante non è il terminus ma l'apertura di un nuovo processo orientato verso la tonica. **Aktionsart: apertura di accomplishment.**

Misure 13-18 (chiusura sezione B): progressione che ritorna alla tonica attraverso un semicadenza (misura 14), una zona di attività modulante (misure 15-16), e l'accomplishment conclusivo che si realizza nella PAC di misura 18. **Aktionsart: accomplishment di chiusura con achievement terminale.**

Fase 3: configurazione di diatesi

Poiché il brano è essenzialmente monodico — una sola linea armonica alla mano sinistra e una melodia alla mano destra che si muovono insieme — la diatesi è quasi sempre sincrona: non ci sono due voci in Aktionsart diverse. Tuttavia, in certi momenti la mano sinistra e la mano destra mostrano una leggera dissincronia aspettuale: alla misura 7, per esempio, la mano sinistra è già nell'accomplishment cadenzale (la progressione ii6-V7 è avviata) mentre la mano destra è ancora nell'ornamentazione atelica della melodia. Questa microsincronia — durata meno di una misura — produce la qualità di "sospensione" che si sente subito prima della cadenza della misura 8.

Fase 4: curva di tensione aspettuale del tema

La curva di tensione aspettuale del tema ha la seguente forma: bassa (misure 1-4, stato), progressivamente crescente (misure 5-7, transizione attività-accomplishment), massima locale (misura 7.5-8, punto più vicino al terminus), riduzione post-achievement (misura 8), nuovo aumento (misure 9-12, apertura di nuovo accomplishment), tensione sostenuta (misure 13-16), massima dell'intero tema (misura 17), risoluzione finale (misura 18).

Confrontata con la curva di tensione armonica di Lerdahl (che cresce sull'armonia di dominante e si risolve sulla tonica), la curva aspettuale è quasi sovrapponibile — nel tema, la grammaticalizzazione della telicità del classicismo viennese è così efficace da rendere le due curve quasi identiche. La divergenza emerge nelle variazioni.

Fase 5: le variazioni come trasformazioni aspettuative

La prima variazione (misure 19-36) ha la stessa struttura armonica del tema ma una figurazione più rapida di sedicesimi. Aspettualmente, la figurazione rapida produce una coercion telicizzante: le figure di sedicesimi sono ateliche prese da sole (potrebbero continuare indefinitamente), ma nel contesto della struttura cadenzale del

tema acquisiscono direzionalità, vengono percepite come accomplishments locali che si susseguono rapidamente. La curva di tensione aspettuale della prima variazione è più "granulare" di quella del tema: ci sono più momenti di micro-telicità locale, più piccoli achievements, una densità di eventi aspettuati molto più alta.

La terza variazione (in la minore) trasforma radicalmente la struttura aspettuale. Il modo minore produce una qualità stativa più marcata nelle zone di presentazione: la stessa struttura armonica, in minore, produce stati che hanno un carattere di "peso" e di persistenza che in maggiore non erano presenti. Gli accomplishments cadenzali hanno in minore una qualità di inevitabilità più cupa: il terminus è lo stesso (la tonica) ma il processo verso di esso ha una direzionalità più carica, più tesa. Aspettualmente, la terza variazione ha una proporzione maggiore di eventi stativi e una qualità più intensa degli accomplishments: è il caso di una trasformazione della struttura aspettuale attraverso il modo, indipendentemente dalla struttura armonica.

La sesta variazione (Adagio) è il caso più interessante. La figurazione lentissima e ornamentale, con abbellimenti, trilli e valori puntati, produce una strana qualità aspettuale: la struttura cadenzale è la stessa del tema, ma la lentezza del tempo rallenta così tanto il dispiegarsi del processo verso il terminus che le fasi dell'accomplishment diventano individualmente percepibili. Si percepisce il "cammino verso" più che il "terminus": l'accomplishment rallenta fino a diventare quasi stativo nel mezzo del suo percorso. Questa trasformazione aspettuale attraverso il tempo è la più sottile della serie e corrisponde, in termini linguistici, all'effetto dell'aspetto progressivo imperfettivo: si è "nel mezzo di un evento" in modo così prolungato che l'orientazione verso il terminus recede in secondo piano rispetto alla presenza nel processo.

Fase 6: narrativa aspettuale del movimento intero

Il movimento intero segue una narrativa aspettuale che si può descrivere come: progressiva complessificazione seguita da semplificazione e recupero. Il tema presenta la struttura aspettuale nella sua forma più semplice e trasparente. Le prime due variazioni aumentano la densità degli eventi aspettuati (più micro-achievements, più transizioni rapide). La terza variazione (minore) introduce una variante qualitativa che modifica il carattere degli stati e degli accomplishments senza cambiarne la struttura. Le variazioni quarta e quinta aumentano progressivamente la densità fino al culmine energetico. La sesta variazione (Adagio) realizza la semplificazione finale: la struttura aspettuale del tema torna, rallentata e ornamentata, come distillazione aspettuale di tutto il percorso. L'ultima cadenza della variazione sei è l'achievement finale del movimento intero — il terminus dell'accomplishment di macro-scala che il tema aveva aperto all'inizio.

IV. Le neuroscienze del ritmo musicale: una base empirica per la teoria aspettuale

Il problema del "tempo interno" nella neuroscienza

La teoria aspettuale proposta in questa monografia postula che gli ascoltatori percepiscano eventi musicali come aventi strutture temporali interne qualitativamente diverse — stati, attività, accomplishments, achievements — e che questa percezione produca risposte cognitive distinte. Questa postulazione, per quanto teoricamente motivata, richiederebbe una base neuroscientifica per essere pienamente convincente. Gli ultimi vent'anni di ricerca neuroscientifica sulla percezione del tempo musicale hanno prodotto risultati che permettono di iniziare la costruzione di questa base, anche se nessuna ricerca ha ancora esplicitamente mirato a testare le categorie aspettuative nel senso di questa monografia.

I risultati più rilevanti provengono da tre linee di ricerca convergenti: la ricerca sul *neural entrainment* (trascinamento neurale nei ritmi musicali), la ricerca sulle risposte cerebrali alle violazioni di aspettativa armonica e metrica, e la ricerca di Aniruddh Patel sul rapporto tra struttura ritmica linguistica e struttura ritmica musicale.

Neural entrainment e la struttura gerarchica del tempo

La ricerca sul neural entrainment ha mostrato che il cervello "si sincronizza" ai ritmi musicali attraverso oscillazioni neurali che seguono la periodicità della musica — un risultato dimostrato per la prima volta in modo sistematico da Nozaradan, Peretz, Mouraux e Vuust in una serie di lavori dal 2011 al 2016. La scoperta cruciale è che questo entrainment non si limita alla frequenza fondamentale del ritmo: il cervello si sincronizza simultaneamente con i multipli della frequenza fondamentale, producendo una rappresentazione neurale della struttura gerarchica del metro musicale. Una misura di quattro quarti produce entrainment alla frequenza della semiminima, ma anche alla frequenza della minima (gruppi di due), della semibreve (la misura intera) e della doppia semibreve (il gruppo di due misure).

Questa struttura gerarchica del neural entrainment ha un'implicazione diretta per la teoria aspettuale: essa suggerisce che il cervello costruisce automaticamente rappresentazioni temporali a scale diverse, esattamente come la teoria aspettuale postula che le Aktionsart si applicano a livelli diversi di analisi (micro, meso, macro). La diatesi aspettuale — la coesistenza di diverse strutture temporali nelle diverse voci — ha dunque un correlato neurale plausibile: il cervello, che rappresenta già la struttura metrica a scale multiple, potrebbe facilmente estendere questa

rappresentazione alle strutture aspettuali delle singole voci, tracciando la struttura di ognuna a livelli diversi della gerarchia temporale.

Le risposte ERN e ERAN alle violazioni di aspettativa

La ricerca neuroscientifica sull'elaborazione armonica ha identificato due risposte neurali alle violazioni di aspettativa: l'*Early Right Anterior Negativity* (ERAN) per le violazioni di aspettativa armonica, e la *Early Negativity* (EN) per le violazioni di aspettativa metrica. Entrambe le risposte hanno latenza breve (100-200 millisecondi) e si distinguono chiaramente nel segnale EEG, suggerendo che il cervello processa le violazioni di aspettativa armonica e metrica attraverso meccanismi distinti.

La domanda rilevante per la teoria aspettuale è: esiste una risposta neurale specifica alle violazioni di aspettativa *aspettuale*? In altre parole, se si interrompe un accomplishment cadenzale prima del terminus (come quando il pianista si ferma nel mezzo di un'esposizione sonatistica senza eseguire la cadenza), si osserva una risposta neurale distinta da quella prodotta dalla violazione armonica (un accordo dissonante inaspettato) e dalla violazione metrica (un cambio di metro improvviso)? La ricerca esistente non ha ancora risposto direttamente a questa domanda, ma le premesse teoriche e metodologiche per farlo esistono, e il suo disegno sperimentale è relativamente diretto.

La PRIMIR hypothesis di Patel: ritmo linguistico e ritmo musicale

Aniruddh Patel, nel suo *Music, Language, and the Brain* (2008), ha proposto la *PRIMIR hypothesis* (Processing Rhythm In Music Is Related) secondo cui il ritmo musicale e il ritmo linguistico condividono risorse cognitive neurali e che le differenze tra le strutture ritmiche delle lingue si riflettono nelle preferenze ritmiche delle musiche di quelle stesse culture. La tesi è supportata da una serie di studi cross-linguistici che mostrano come i parlanti di lingue con diversa struttura ritmica (stress-timed vs. syllable-timed) abbiano diverse sensibilità alle strutture ritmiche musicali.

Questa ipotesi è direttamente rilevante per la teoria aspettuale in quanto — se corretta — suggerisce che il sistema cognitivo per il trattamento del ritmo è condiviso tra linguaggio e musica, esattamente come la teoria aspettuale propone che le strutture temporali interne degli eventi (le Aktionsart) siano condivise tra dominio linguistico e dominio musicale. In effetti, la PRIMIR hypothesis di Patel può essere vista come una conferma empirica parziale dell'ipotesi più ampia di questa monografia: se la struttura ritmica è cognitivamente condivisa tra musica e linguaggio, è plausibile che anche la struttura aspettuale — che è la controparte semantica della struttura ritmica, la descrizione delle proprietà temporali interne degli eventi piuttosto che del loro ritmo esterno — sia condivisa tra i due domini.

Kotz e Schwartze: il modello cortico-sottocorticale del timing

Sonja Kotz e Michael Schwartze hanno proposto il *BUMP model* (*Basal ganglia, Uvula, Motor cortex, Prefrontal cortex*) per descrivere il circuito neurale che governa la previsione temporale e il timing nella musica e nel linguaggio. Il modello sostiene che i gangli della base — struttura sottocorticale associata tradizionalmente al controllo motorio — svolgono un ruolo centrale nella previsione di eventi temporali: il cervello usa i gangli della base per costruire un modello del ritmo in corso e prevedere quando arriveranno gli eventi futuri, sia nel parlato che nella musica.

La rilevanza per la teoria aspettuale è indiretta ma significativa. Il BUMP model implica che la previsione temporale sia un meccanismo fondamentale della percezione musicale e linguistica, e che questo meccanismo operi a livello sottocorticale, "prima" dell'elaborazione consapevole e semantica. Se la struttura aspettuale degli eventi musicali — la direzionalità degli accomplishments, la stabilità degli stati, la puntualità degli achievements — produce risposte di previsione temporale sistematicamente diverse, queste differenze dovrebbero essere visibili nei correlati neurali prodotti dai gangli della base, anche indipendentemente dall'elaborazione corticale consapevole.

Zatorre e il sistema del "reward" musicale

Robert Zatorre e colleghi hanno mostrato, attraverso studi di neuroimaging, che la musica attiva il sistema dopaminergico del reward — lo stesso sistema attivato dal cibo, dal sesso e dalle droghe — in momenti specifici dell'ascolto: i cosiddetti "brividi musicali" (*chills* o *frissons*) si verificano preferibilmente nei momenti di massima aspettativa soddisfatta o violata. Il picco di attività dopaminergica precede leggermente il momento di massima tensione armonica e poi si scarica nel momento della risoluzione.

Tradotto in termini aspettuali: il picco di attività dopaminergica si concentra nei momenti di massima tensione verso il terminus dell'accomplishment e nel momento del suo achievement terminale. La "ricompensa" neurale della musica è, almeno parzialmente, la ricompensa dell'accomplishment completato — della struttura telica che raggiunge il suo terminus. Questa interpretazione suggerisce che la struttura aspettuale degli eventi musicali non è solo cognitivamente rilevante ma affettivamente saliente: la distinzione tra telico e atelico, tra accomplishment e attività, tra processo verso terminus e processo senza terminus, è una distinzione che il sistema nervoso tratta come affettivamente significativa.

Verso una neuroscienza dell'aspetto musicale: un programma di ricerca

I risultati sopra descritti non costituiscono ancora una neuroscienza dell'aspetto musicale, ma indicano i contorni di un programma di ricerca plausibile. Il passo successivo più immediato sarebbe un esperimento di neuroimaging (fMRI o EEG ad alta densità) che presenti agli ascoltatori brani musicali sistematicamente manipolati nella loro struttura aspettuale — brani in cui la stessa struttura armonica è realizzata

con Aktionsart diverse attraverso modifiche di tempo, ritmo e ornamentazione — e misuri le differenze nelle risposte neurali. Se la teoria aspettuale è corretta, si dovrebbero osservare differenze sistematiche nelle risposte corticali e sottocorticali che non sono spiegabili dalle sole differenze armoniche o ritmiche. Un secondo passo sarebbe uno studio cross-linguistico che confronti ascoltatori di lingue con aspetto obbligatorio (russo, cinese) con ascoltatori di lingue senza aspetto obbligatorio (swahili, turco) nella sensibilità alle strutture aspettuative musicali: se la struttura aspettuale è cognitivamente condivisa tra musica e linguaggio, ci si aspetterebbe che i parlanti di lingue con aspetto obbligatorio mostrino una sensibilità diversa — non necessariamente maggiore ma qualitativamente differente — alle strutture aspettuative musicali rispetto ai parlanti di lingue senza aspetto. Questo programma di ricerca è tecnicamente realizzabile con metodologie esistenti, e la sua realizzazione trasformerebbe la teoria aspettuale della musica da proposta teorica speculativa a ipotesi empiricamente verificata.

Bibliografia Ragionata

Opere fondamentali sull'aspetto linguistico

Vendler, Zeno (1957). "Verbs and Times." *The Philosophical Review*, 66(2), pp. 143-160. L'articolo fondativo della tassonomia aspettuale in quattro classi. Nonostante la sua età, rimane il riferimento imprescindibile per la distinzione delle Aktionsart, e la semplicità dei test operazionali proposti — compatibilità con il progressivo, risposta alle domande "per quanto tempo?" versus "in quanto tempo?", logica dell'interruzione — rende il sistema applicabile anche al dominio musicale senza necessità di modifiche sostanziali. Ripubblicato in *Linguistics in Philosophy* (Cornell University Press, 1967).

Comrie, Bernard (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge University Press. La sistematizzazione crosslinguistica più influente della teoria aspettuale. Introduce la distinzione perfettivo/imperfettivo come categoria grammaticale distinta dalle Aktionsart, e analizza il progressivo, l'abituale, il risultativo e l'esperienziale come sottocategorie dell'imperfettivo. Fondamentale per la prospettiva tipologica del Capitolo III.

Dowty, David R. (1979). *Word Meaning and Montague Grammar*. Reidel. La formalizzazione logica più rigorosa delle Aktionsart di Vendler, con l'introduzione

dell'operatore BECOME per descrivere i cambiamenti di stato degli achievement. Tecnico ma necessario per chi voglia costruire una formalizzazione più rigorosa di quella sviluppata nell'Appendice A.

Krifka, Manfred (1989). "Nominal Reference, Temporal Constitution, and Quantification in Event Semantics." In Renate Bartsch, Johan van Benthem e Peter van Emde Boas (a cura di), *Semantics and Contextual Expression*. Foris, pp. 75-115. L'introduzione della nozione di "cumulatività" e "quantizzazione" come proprietà formali che distinguono attività da accomplishments. Il criterio della cumulatività usato nel Capitolo II è derivato direttamente da questo lavoro.

Krifka, Manfred (1992). "Thematic Relations as Links between Nominal Reference and Temporal Constitution." In Ivan Sag e Anna Szabolcsi (a cura di), *Lexical Matters*. CSLI Publications, pp. 29-53. Sviluppo e approfondimento del sistema del 1989, con particolare attenzione alla relazione tra struttura eventiva e struttura argomentale.

Moens, Marc e Steedman, Mark (1988). "Temporal Ontology and Temporal Reference." *Computational Linguistics*, 14(2), pp. 15-28. L'introduzione della nozione di "nucleo eventivo" — preparazione, culminazione, stato conseguente — e del meccanismo di coercion aspettuale. Il meccanismo di coercion sviluppato nel Capitolo IV riprende direttamente questo lavoro, che resta uno dei contributi più influenti nella semantica computazionale del tempo.

Brinton, Laurel J. (1988). *The Development of English Aspectual Systems*. Cambridge University Press. L'estensione dell'analisi aspettuale al di là dei verbi, con applicazione ai predicati nominali e aggettivali. Rilevante per la proposta che la struttura aspettuale musicale si applichi a qualsiasi struttura sonora con struttura temporale interna.

Smith, Carlota (1991). *The Parameter of Aspect*. Kluwer. Una trattazione completa e sistematica dell'aspetto nelle lingue naturali, con ampia prospettiva crosslinguistica. Particolarmente utile per la distinzione tra aspetto e Aktionsart e per la discussione delle lingue senza morfologia aspettuale obbligatoria.

Pustejovsky, James (1991). "The Syntax of Event Structure." *Cognition*, 41, pp. 47-81. La proposta della struttura qualia come rappresentazione dell'Aktionsart come parte del significato lessicale. Rilevante per la discussione della relazione tra struttura eventiva e struttura semantica in contesti di ambiguità aspettuale.

Teoria musicale

Caplin, William E. (1998). *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*. Oxford University Press. La teoria della forma classica più sistematica disponibile. La distinzione tra funzioni formali "tight-knit" e "loose" è una teoria implicita della grammaticalizzazione aspettuale nel

classicismo viennese, come mostra l'analisi del Capitolo III. Fondamentale per chiunque si occupi di aspetto nella musica tonale.

Hasty, Christopher (1997). *Meter as Rhythm*. Oxford University Press. La proposta più radicale e più originale di una teoria del metro come processo temporale emergente. Il Capitolo IV di questo saggio rilegge Hasty attraverso la lente aspettale, mostrando che la sua teoria del "becoming" metrico è implicitamente una teoria dell'aspetto ritmico. Lettura lenta e difficile, ma indispensabile per pensare seriamente il tempo musicale.

Jackendoff, Ray (1991). "Musical Parsing and Musical Affect." *Music Perception*, 9(2), pp. 199-229. La proposta che le strutture concettuali del linguaggio — incluse le strutture eventive che sottostanno all'aspetto — siano applicabili alla musica attraverso l'"isomorfismo tematico". Meno sistematico di Jackendoff (1983), ma più direttamente rilevante per il problema aspettale.

Jackendoff, Ray (1983). *Semantics and Cognition*. MIT Press. Il quadro teorico fondamentale in cui Jackendoff sviluppa la teoria delle funzioni concettuali (GO, BE, CAUSE, CHANGE) e la loro applicazione a domini diversi. Necessario per comprendere le implicazioni più ampie della proposta sviluppata nel paragrafo 1.5.

Lerdahl, Fred e Jackendoff, Ray (1983). *A Generative Theory of Tonal Music*. MIT Press. La teoria gerarchica della struttura musicale tonale più influente. La teoria della riduzione metrica e gruppale fornisce la base per la segmentazione degli eventi musicali in unità di analisi aspettale, come discusso nell'Appendice A.

Lerdahl, Fred (2001). *Tonal Pitch Space*. Oxford University Press. La formalizzazione dello spazio tonale come spazio cognitivo con geometria determinata. La curva di tensione armonica di Lerdahl, confrontata con la curva di tensione aspettale proposta in questo saggio, permette di distinguere le due dimensioni della tensione musicale, come proposto nel Capitolo IV.

Koch, Heinrich Christoph (1793). *Versuch einer Anleitung zur Composition*, vol. 3. Böhme. La prima sistematizzazione della gerarchia cadenzale nella teoria della forma classica, citata come fonte storica della grammatica della telicità graduata discussa nel Capitolo III.

Ratz, Erwin (1951). *Einführung in die musikalische Formenlehre*. Österreichischer Bundesverlag. La teoria della forma musicale che introduce la distinzione tra Hauptsatz e Nebensatz come funzioni formali con diverso grado di telicità. Precursore diretto di Caplin.

Schenker, Heinrich (1935). *Der freie Satz*. Universal Edition. La teoria della struttura di fondo della musica tonale, citata nel Capitolo II a proposito dell'obiezione che tutta la musica tonale è telica. La risposta a questa obiezione richiede di distinguere il livello

della struttura di sistema (descritto da Schenker) dal livello della struttura cognitiva dell'esperienza (descritto dalla teoria aspettuale).

Psicologia cognitiva della musica

Bharucha, Jamshed J. e Stoeckig, Kristi (1986). "Reaction Time and Musical Expectancy: Priming of Chords." *Journal of Experimental Psychology: Human Perception and Performance*, 12(4), pp. 403-410. La dimostrazione sperimentale che l'aspettativa armonica ha effetti misurabili sui tempi di elaborazione. Fornisce la base empirica per la proposta che l'aspetto musicale sia cognitivamente reale e non semplicemente una proiezione teorica.

Bigand, Emmanuel, Parncutt, Richard e Lerdahl, Fred (1996). "Perception of Musical Tension in Short Chord Sequences." *Perception & Psychophysics*, 58(1), pp. 125-141. La verifica sperimentale che la tensione armonica percepita correla con la distanza nel tonal pitch space di Lerdahl, fornendo la base empirica per la distinzione tra tensione armonica e tensione aspettuale discussa nel Capitolo IV.

Huron, David (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press. La teoria dell'aspettativa musicale più sistematica e meglio supportata empiricamente. Le sue categorie — immaginazione, tensione, predizione, reazione, valutazione — si relazionano con le categorie aspettuali in modi che questo saggio ha solo parzialmente esplorato.

Kotz, Sonja A. e Schwartz, Michael (2010). "Cortical-subcortical circuits in speech-timing: An introduction to the BUMP model." *Brain & Language*, 115(1), pp. 80-88. Il modello neurofisiologico del timing musicale e linguistico, rilevante per il collegamento tra struttura aspettuale e meccanismi neurali del tempo.

Semiotica e narratologia musicale

Tarasti, Eero (1994). *A Theory of Musical Semiotics*. Indiana University Press. La proposta di una semiotica musicale ispirata a Greimas, discussa nelle Conclusioni in relazione alla teoria aspettuale. La distinzione tra isotopie musicali e Aktionsart musicali è produttiva anche quando i due approcci divergono nella loro impostazione di fondo.

Nattiez, Jean-Jacques (1987). *Musicologie générale et sémiologie*. Bourgois. Il quadro tripartito (poietico, esthésique, neutro) per l'analisi del significato musicale. La collocazione della teoria aspettuale nel livello esthésique, discussa nelle Conclusioni, è necessaria per situare questo saggio nella tradizione della semiotica musicale.

Zbikowski, Lawrence M. (2002). *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. Oxford University Press. La proposta di una comprensione musicale basata su strutture concettuali metaforiche. Il confronto tra questa proposta e la teoria asettuale illumina la differenza tra "comprensione per metafora" e "struttura cognitiva condivisa".

Etnomusicologia e tradizioni non-occidentali

Arom, Simha (1985). *Polyphonies et polyrythmies instrumentales d'Afrique centrale*. SELAF. La ricerca più sistematica sulle strutture polifoniche africane, con particolare attenzione al meccanismo dell'interlocking. Fondamentale per l'analisi sviluppata nel Capitolo V.

Berliner, Paul (1994). *Thinking in Jazz: The Infinite Art of Improvisation*. University of Chicago Press. La ricerca etnografica sull'improvvisazione jazz che fornisce la base per la comprensione del bebop e del jazz modale come sistemi asettuali, attraverso la descrizione delle pratiche cognitive degli improvvisatori.

Clayton, Martin (2000). *Time in Indian Music: Rhythm, Metre, and Form in North Indian Rāg Performance*. Oxford University Press. L'analisi più sistematica del sistema del tala nella prospettiva della teoria cognitiva del ritmo. Fondamentale per l'analisi del Capitolo V, in particolare per la discussione del sam come terminus asettuale ciclico.

Hood, Mantle (1954). *The Nuclear Theme as a Determinant of Patet in Javanese Music*. Wolters. Il lavoro fondativo dell'analisi del gamelan giavanese nella tradizione musicologica occidentale. La nozione di "nuclear theme" e la sua relazione con la struttura colotomica sono discusse nel Capitolo V.

Perlman, Marc (2004). *Unplayed Melodies: Javanese Gamelan and the Genesis of Music Theory*. University of California Press. Un'analisi sofisticata del gamelan giavanese che mette in discussione alcune semplificazioni della musicologia tradizionale sul sistema tonal del gamelan. Particolarmente rilevante per la cautela metodologica espressa nel paragrafo 5.6.

Potter, Keith (2000). *Four Musical Minimalists: La Monte Young, Terry Riley, Steve Reich, Philip Glass*. Cambridge University Press. La trattazione più sistematica del minimalismo americano da una prospettiva musicologica. Fondamentale per l'analisi del paragrafo 3.5.

Shankar, Ravi (1968). *My Music, My Life*. Simon & Schuster. L'autobiografia del più celebre musicista di raga nordindiano nel Novecento, con ampia discussione del sistema del tala da una prospettiva interna alla tradizione. Complementare alle analisi teoriche di Clayton e Rowell.

Rowell, Lewis (1992). *Music and Musical Thought in Early India*. University of Chicago Press. Una trattazione storica e teorica della musica indiana e del suo sistema concettuale, con discussione approfondita della nozione di ciclo temporale e del suo rapporto con la struttura formale dell'improvvisazione sul raga.